Wireless LAN Service Quality Optimization in Academic Environments

Moses E. Ekpenyong¹, Uduak A. Umoh¹, Udoinyang G. Inyang¹ & Aniekpeno M. Jackson¹
¹Department of Computer Science, University of Uyo, P.M.B. 1017, 520003, Uyo, Akwa Ibom State, Nigeria
Correspondence: Moses E. Ekpenyong, Department of Computer Science, University of Uyo, Nigeria. E-mail: mosesekpenyong@uniuyo.edu.ng, mosesekpenyong@gmail.com

Received: August 30, 2017      Accepted: September 19, 2017      Online Published: September 30, 2017
doi:10.5539/mas.v11n10p166             URL: https://doi.org/10.5539/mas.v11n10p166

The research is financed by The Tertiary Education Trust Fund (TETFund) National Research Fund (NRF) Grant (Ref. No. TETFUND/NRF/UNI/UYO/STI/VOL.1/BE)

Abstract
This paper targets optimized service quality (SQ) – a metric that compares the perceived performance by users with the expected performance – sufficient to satisfy users’ quality of experience (QoE). The perceived performance was obtained in a field survey from an academic environment, and using Interval Type-2 Fuzzy Logic (IT2FL), uncertainties inherent in the field data were efficiently modeled for accurate estimation of the SQ. To obtain the expected performance, two unsupervised tools: the Principal Component Analysis (PCA) and Self-organizing Map (SOM) were exploited to abstract the most relevant features, and observe similarity patterns between the abstract features. An Adaptive Neuro-Fuzzy Inference System (ANFIS) was then used to optimize the system performance. Results obtained showed that ANFIS sufficiently optimized and modeled the SQ – as the root mean square error (RMSE) values of the train and test data were approximately the same – for all the study sites considered. However, combining the three campuses produced the least mean absolute error (MAE) of 0.0979 for train data, and the highest MAE of 0.7345 for test data. Further, the least MAE of 0.4707 for test data was obtained from town campus Annex. The wide variation in MAE observed in the train and test data might not be unconnected with the high degree of uncertainties associated with interference, site topology and terrain issues – exhibited by the system under study, as well as the quality of data collected. The proposed system framework has the potentials to develop into a complete location-based system.

Keywords: Location-aware system, neuro-fuzzy approach, principal component, service optimization, self-organizing map, wireless LAN

1. Introduction
The deployment and utilization of wireless local area network (WLAN) technology in academic environments are gaining wide acceptance, but they can also impose great challenges that are significantly different from similar deployments at enterprise or commercial settings. The key challenges of scaling a campus WLAN arise from domain-specific issues namely: users’ mobility at different locations within the campus, users’ density, devices, and instantaneous loads caused during peak periods. Hence, the strengths of radio frequency (RF) signals arriving from various APs in a WLAN are related to the position of the user and are vital for deriving the user’s location. As such, knowledge of location and robust models are indeed necessary to minimize cognitive load on users in context- and location-aware systems (Djuknic and Richton, 2001; Anhalt, Smailagic, Siewiorec, Gemperle, Salber, Weber, Beck and Jennings, 2001). The goal of a location-aware system is to convey orientation and search reference necessary for cognitively ergonomic communication, and contributes to the interface design. To actualize true ubiquitousness, next generation location-aware systems are most likely to exploit mobile platforms that integrate multiple sensors and measurement systems – to enable continuous location scanning, graphic visualizations and platform orientation (Kealy, Winter and Retscher, 2007). To communicate with users in a more cognitive manner, location aware services (LASs) require techniques for translating position into location information (a reference and abstract location within the mobile platform). In real-time ubiquitous systems, the choice of appropriate locations and positioning requires absolute intelligence to ensure the synchronization of positioning and location. Figure 1 demonstrates the interaction/feedback cycle.
between positioning and localization.

The localization component assess spatial data about the environment including (i) structures, places, relations – for the accurate extraction of users position, (ii) deployed infrastructure and other measurement context; while the positioning component gathers information about spatial context of the current environment, and is useful for specifying the required accuracy, measurement design specification, or constrained position estimation. Location estimation constitutes an important component of location-aware applications – as positioning based on the received signal strength (RSS) is considered a promising and cost effective solution (Kealy, Winter and Retscher, 2007).

Location-based service (LBS) is often considered as part of or a subset of context-aware service, from where location-aware service (LAS) takes its root (Kupper, 2005). A system is context-aware if it exploits context to provide relevant information and/or services to the end-user – where relevancy depends on the user’s task, and information or service are triggered from context information (parameters within the user’s environment and relevant to its task). These parameters may be subdivided into personal, technical, spatial, social, and physical contexts. The major problem associated with context-aware services is how to model these parameters in a quantifiable and computable way. Intelligent location-aware systems exploit state-of-the-art tools such as machine learning and neural computing to minimize error levels associated with user’s location determination. They provide effective training, classification and visualization of salient system features – to manage poor service quality in mission critical systems, which require continuous system calibration and updates of clients’ services – as well as uninterruptible access to available resources.

This paper therefore proposes an intelligent location-aware system for WLAN service quality optimization in academic environments. Locations of the WLAN users are estimated based on extensive radio signal strength measurements – through field trial survey. The problem is regarded as a machine learning one, which models observed signal strength patterns across the study area. Further, the position of clients/users within an AP is controlled by the respective Network Operating Centers (NOCs) in the study area. As such, our design takes full advantage of the numerous access points to improve accuracy and ensure autonomous self-reconfiguration with minimal operator participation. To impose the required geographical boundaries, a survey of each client location (in our case offices/buildings with installed access point) within the coverage area was carried out through extensive measurements in real settings – of mobile and fixed access points infrastructure that received and transmit signal strength from each other. The obtained data (information about the physical location of clients) were then trained and classified to leverage on the spot resolution of poor service quality in the system. This activity demonstrates the capability of our framework to self-reconfigure or auto-calibrate the system, and guarantee accurate services while ensuring network scalability to the growing number of access points.

2. Significance of the Study

A myriad of challenges usually accompany WLAN design and implementation. The notable ones being security vulnerabilities, radio signal interference, multi-path propagation, roaming issues, battery limitations, interoperability problems, and installation issues. These challenges occur because as Wireless LANs propagate data over buildings, campuses, and even cities, the radio signals often travel beyond the limits of the service area – beyond the control of the organization. For instance, radio waves seamlessly penetrate building walls and can be received from any distance within the coverage limits of the network. Further, the effect of distance on signal strength attenuation largely depends on the environmental characteristics and is strongly influenced by construction materials, walls disposition, system capacity, and movement patterns generated by objects in the propagation environment. To address these challenges, a continuous calibration process through periodic computation of new values to accommodate dynamic environmental changes without service disruptions is required, instead of pre-establishing model parameters. Hence, this paper has the potentials to offer the
following:

Autonomous calibration: Autonomous system calibration is required periodically to gather signal attenuation data within the service area, and is necessary to pre-establish the right parameters for building a suitable pathloss model – a mathematical function for predicting with reasonable accuracy, signal attenuation within the service area. In this study, we obtained propagation characteristics from a deployed WLAN/WiFi environment – useful for intelligent classification and prediction of WLAN SQ in academic environments. Although the current data capturing procedure is heavily supervised, and employed the use of the global positioning system (GPS) device – to capture location information; and a WiFi analyzer – to capture relative signal strength indication (RSSI) information; a continuous (unsupervised) calibration of the service area is necessary – using sensors – and is useful for real-time monitoring, detection and resolution of the poor service quality. In this paper, the unsupervised calibration is achieved using the self-organizing map (SOM).

Access control: The proposed framework shall improve access control by allowing network administrators to define geographic boundaries for wireless services and minimize undesired interference that leads to poor service quality (SQ). This paper therefore exploits inherent properties of the wireless channel, and captures client transmissions through signal strength measurements gathered at different service locations, as they use the network. The collected data are then refined to obtain accurate estimates of the service quality at the affected locations.

Accountability: Our proposed framework will optimize accountability – as the use of localization systems is essential for real-time information management – about the physical location of transmitters (client APs, NOCs, etc.), and relies on clients information for proper accountability of users. To distinguish between the devices in the midst of Medium Access Control (MAC) address spoofing, each transmitter/AP is assigned a signal strength pattern or signalprint for proper location-based device identification/representation using the Interval Type-2 Fuzzy Logic (IT2FL) system. This process guarantees that all devices are properly accounted for and monitored.

Scalability: The current design can be scaled to accommodate more access points without incurring undue costs penalties and overdependence on physical reconfiguration. Although research works suggest that centralized architectures achieve lower costs, the proposed system expects existing (deployed) access points to reconfigure themselves without physical intervention. Hence, information regarding poor SQ within the coverage area required by the localization mechanisms can be autonomously established by the system with minimal impact on accuracy.

3. Related Works

Numerous systems and emerging technologies have been proposed in the literature to determine the location of users for mobile computing applications – with the GPS – a satellite-based navigation aid – originally developed by the US military, being the pioneering one. GPS has been very successful in open areas, but ineffective in indoor settings or in urban areas with tall obstacles such as buildings shielding the satellite signals. Other approaches to location based services include: Microsoft research RADAR location system (Paramvir, Venkata and Padmanabhan, 2000), where the radio frequency (RF) signal strength is used as a measure of distance between the access point (AP) and mobile terminal. In Roberto, Thang and Alessandro (1999), a method to determine the locations of mobile terminal in high-speed wireless LAN environment using the IEEE 802.11b standard was proposed. Their methodology explored the use of neural network models and automated learning techniques. As is the case for the RADAR system, no additional special-purpose equipment was required. While the flexible modeling and learning capabilities of neural networks achieved lower errors in position determination, they were amenable to incremental improvements.

Location-aware computing is an interesting area of research that exploits the potentials of modern communication technology (Brown, Bovey and Chen, 1997; Chen and Kotz, 2000; Hightower and Borriello, 2001; Leonhardt, 1998). The location of the mobile terminal can be estimated using radio signals transmitted or received by the terminal (Rappaport, Reed, and Woerner, 1996; Bahl and Padmanabhan, 2000; Bulusu, Heidemann and Estrin, 2000; Myllymaki, Roos, Tirir, Misikangas, Sievanen, 2002; Castro, Chiu, Kremenek, and Muntz, 2001; Syrjarinne, 2001; Youssef and Agrawala, 2002). In Myllymaki, Roos, Tirir, Misikangas and Sievanen (2002), the location of a WLAN user based on radio signal strength measurements was performed using the user’s mobile terminal. In their approach, the location estimation was considered as a machine learning problem, and with a probabilistic framework, the location estimation problem was solved. The feasibility of their approach was also demonstrated by reporting results from field tests – where a probabilistic location estimation method was validated in a real-world indoor environment.

Chen, Huang and Song (2009) discussed the problem of user location estimation using the received signal
strength in a radio-frequency wireless local area network (WLAN). They proposed a machine learning framework and employed a multivariable Gaussian distribution model with an offline training phase plus a real-time location phase. In the offline phase, they recorded the received signal strength of numerous training locations and computed the parameters of the Gaussian distribution; while in the real-time phase user’s location was determined by matching the received signal strength patterns against the training patterns. Their experiments demonstrated that the proposed method had lower location errors.

Sharma and Singh (2010) after creating spatial database from environmental information about a study area trained and mapped location-independent parameters to a specific area of interest – to establish the probable location of users within the network. The target tradeoffs investigated were platform selection, node deployment and network. In Farid, Nordin, Ismail and Abdullah (2016), a hybrid technique for implementing indoor localization that adopts fingerprinting approaches in both WiFi and Wireless Sensor Networks (WSNs) was proposed. Their model exploited machine learning technique, precisely: artificial neural network (ANN), for position computation. Experimental results showed that the proposed hybrid system improved accuracy by reducing the average distance error, and applying genetic algorithm (GA) based optimization technique did not improve the accuracy, further. In Wang, Shi and Wu (2017), a Gaussian filtering algorithm based on an extreme learning machine (ELM) was proposed to address the problem of inaccurate indoor positioning in the midst of significant RSSI fluctuations during the measurement process. The proposed positioning system was tested in a real experimental environment, and found to achieve higher position accuracy and speedup, compared to previous algorithms.

4. System Framework

The proposed system framework defines the supporting structure of our location-aware service system. As shown in Figure 2, the framework has of a five-layered workflow, namely,

- WLAN test bed design
- Service area definition
- LAS feature taxonomy
- SQ modeling
- SQ optimization

![Figure 2. Proposed LAS framework](image-url)
4.1 WLAN Test Bed Design and Service Area Definition

The Study Environment: Our study environment is the University of Uyo (UNIUYO). UNIUYO has a total land mass of 1,535.055 hectares, and consists of five separate campuses namely, town campus, town campus annex, main campus, University of Uyo Teaching Hospital (UUTH), and Basic Studies Campus. The University deploys an inter-campus WLAN infrastructure that provides communication over a short geographical range using radio signals. The radio signals are propagated using network bridges – to create aggregate networks from either two or more communication networks and/or segments. The propagated signals are then regenerated along the next leg of the transmission medium to overcome the attenuation (loss of signal strength) caused by free-space electromagnetic-field divergence or cable loss, and to extend signals over a distance. The existing infrastructure consists of two layers: the Fibre Optic (FO) layer and the Wireless Network (WN) layer. The FO layer implements the Local Area Network (LAN) infrastructure and connects the various buildings, while the WN layer distributes signals to the buildings. Wireless Access Points (WAPs) are connected to the edge of the fibres to enable clients/users communicate effectively with the Wireless Network Adapters (WNAs). Currently, the three major campuses of the University under study (town campus: covering 56.956 hectares, town campus annex: covering 34.919 hectares, and main campus: covering 1,443.180 hectares) have been fully bridged in an intranet using the FO technology, and are considered in this paper.

The first two phases are accomplished by creating test bed simulators (the WLAN setup environment required for experimenting and testing the validity of the research) to support the visualization interface of the proposed LAS system. The design methodology of the test bed simulators implemented the following steps:

Reconnaissance survey: examining the generic characteristics of the area and establishing the start and end points of the WN and the FO installations within the three campuses.

Database model: integrating a geodatabase model to enable proper planning and management of the WLAN resources. The purpose of this integration is to bootstrap important management plans and provide solutions to users/clients in real time.

Location capture and feature extraction: capturing the various FO milestone and manhole locations using GPS device; and extracting building and road layers within the three campuses from satellite images or base maps of the study area through digitization using the ArcGIS 10.3 software. The extracted features and the reconnaissance survey data formed part of the service area definition, and allowed for the determination of any deviations required in the basic geometric standards to be adopted for the implementation of robust WLAN communication.

Feature mapping: superimposing the GPS data (i.e., manhole and milestone data) on the extracted surface after image digitization to obtain a complete and accurate test bed for the three campuses.

Pathloss distance extraction: simulating the log-distance pathloss for each building from the site-specific data. This feature is important as it aids in the prediction of the attenuation or loss due to distance, experienced by signal inside the building. Since free-space pathloss models are theoretical and not applicable in real-life scenarios, the log-distance path loss model was modified to obtain a practical model for driving the proposed test bed, by introducing additional parameters found in the study area. We also assumed that the received power decreases logarithmically with distance. In theory, the basic log-distance path loss model is generally expressed as,

\[ P_L(dB) = P(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) \]

where,

- \( n \) is the path loss exponent,
- \( d \) is distance between the transmitter and receiver in meters,
- \( d_0 \) is the reference distance, usually 1 kilometre (or 1 mile)
- \( P(d_0) \) is power received at ‘\( d_0 \)’

Now, to optimize the campus WLAN for maximum throughput, the modified log-distance path loss model included the building structure, layout of rooms, and type of construction materials used; and can be rewritten as:

\[ P_L(dB) = I_{env} + 10n\log_{10}(d) + L_f(n_f) + L_w(n_w) + I_m + \sigma \]

where,

- \( I_{env} \) is the indoor environment under study and is dependent on fixed loss factor in dB
\( d \) is the distance between the transmitter and receiver in meters

\( L_f(n_f) \) is the floor penetration loss factor in dB

\( L_w(n_w) \) is the wall penetration loss factor in dB

\( l_m \) is attenuation caused by the indoor material

\( X_\sigma \) is the Normal (Gaussian) random variable in dB, with zero mean and standard deviation \( \sigma \) (dB), with log normal shadowing.

Equation (2) is useful for the computation of indoor pathloss in academic environments and can be simulated to ascertain the loss experienced in the study area. Table 1 catalogues the values of indoor parameters according to environment (Cebula, Ahmad, Graham, Hinds, Wahsheh, Williams and DeLoatch, 2011; Saunders, Kelly, Jones, Dell’Anna and Harrold, 2000).

Table 1. Indoor parameter values for different WLAN environment

<table>
<thead>
<tr>
<th>Environment</th>
<th>Residential</th>
<th>Office</th>
<th>Commercial</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Env(dB) )</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>( 10n )</td>
<td>28</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>( L_f(n_f) + L_w(n_w) )</td>
<td>4n</td>
<td>15 + 4(n − 1)</td>
<td>6 + 3(n − 1)</td>
</tr>
<tr>
<td>( X_\sigma ) (dB)</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Different materials produce varying amount of attenuation in the shadow region. The amount of attenuation is also frequency dependent. Table 2 shows the attenuation caused by major indoor materials used in Nigerian academic environments.

Table 2. Attenuation caused by different indoor materials

<table>
<thead>
<tr>
<th>Material</th>
<th>Cement block (5-9 inches)</th>
<th>Brick</th>
<th>12mm plywood separator</th>
<th>18mm plywood separator</th>
<th>Metal</th>
<th>0.25 inch Glass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss at 2.4 GHz</td>
<td>-3 dB to -4 dB</td>
<td>-4 dB</td>
<td>-0.50 dB</td>
<td>-1.90 dB</td>
<td>-3 dB</td>
<td>-2 dB</td>
</tr>
</tbody>
</table>

4.2 Preliminary Findings

Network Coverage Analysis – Design Issues: In Figure 3, stem plots visualizing the distance of the installed APs from the NOC at the different campuses are presented. We observe that the APs were more separated (from the NOC) at the main campus than at the town campus and town campus annex. This is due to the large land mass and few installed APs to drive the WLAN. The main campus is the permanent campus of the University of Uyo, and is currently under development. This study is therefore apt, as deployment defects and steps ignored at the planning stage in the main campus are most likely to be addressed before full deployment. At the town campus, more APs are located far from the NOC, but clustered in a non-line-of-site fashion due to the dense nature of the environment. This implies that the APs/NOC may not be at the optimal location and certainly require relocation to compensate for the average pathloss due to distance. At the town campus annex, the locations of the APs also vary in distance (from the NOC) but are well separated. Hence, a more comprehensive site survey for automated coverage and signal quality visualization of deployed APs – including the identification of key locations for performance improvement and data communication, is indeed necessary. The benefits of a comprehensive survey include optimization of the existing WLAN infrastructure and RF spectrum analysis. Also, in-depth assessment with complete inventory of all WLAN infrastructure including SQ coverage maps constitute some of the outcomes of this paper. The coverage analysis will ensure real time detection and quick resolution of network coverage issues currently experience in the existing WLAN.
Figure 3. Stem plots of distance of AP from NOC vs. number of some installed APs in the various campuses

**Site Analysis – Topographic Survey:** To contend with the inconsistent topography of the study area, a complete site analysis is necessary to map out appropriate ways of locating BSs and APs, which currently were not professionally done. Hence, predictive and onsite surveys are required to simulate the RF and use the predicted results to verify the operations of the existing WLAN. A passive survey may also be necessary to detect transmission holes at different communication levels/layers. Passive survey involves the collection of RF data from individual APs in the study area to validate the design requirements. In Figure 4, the effect of terrain on installed FO infrastructure at the three major campuses is presented. Figure 4 reveals that the terrain in the town campus annex appears flat, compared with the other two campuses where the terrain looks hilly and depressed at some milestone and manhole points – seemingly close to a ravine constantly threatening a section of the town campus. As such careful planning is expected to reduce the undue variability in SQ delivery to end-users/clients.

Figure 4. Effect of terrain on WLAN installation
4.3 LAS Feature Taxonomy

The proposed LAS taxonomy is a campus wide model with two major components namely operating environment and technological infrastructure. The operating environment includes physical structures relating to spaces (buildings), equipment and tools, as well as end-users. The buildings are associated with classrooms, offices, and laboratories. Other spaces include libraries, residential, social (recreation and relaxation) and commercial spaces. The equipment and tools include teaching, communication, entertainment, information sources, resources and events within and outside of physical environment – where students participate in the learning process directly and remotely (Kuuskorpi and González, 2011). The users include students, lecturers, administrative staff, service providers and guests. The enhanced synergy between students and the teachers, on one hand, and the physical environment, on the other hand, produce the teaching space (c.f., Ellis and Goodyear, 2016). The technological infrastructure consists of information and communications technology related tools for the acquisition, processing, storage and dissemination of information and knowledge between elements of the physical environment. These include hardware, Software, networks, humanware and network operating centre (NOC). Figure 5 documents these features, and the notable ones are discussed in the following subsections.

![Figure 5. Feature taxonomy of our LAS test bed](image-url)
4.3.1 Physical Space

The physical space of the test bed is a multi-location campus structure covering the five campuses within the study area: Main Campus (MC), Town Campus (TC), Town Campus Annex (TCAX), University of Uyo Teaching Hospital (UUTH), and Basic Studies Campus (BSC). These locations are characterized by buildings, equipment/tools, FO infrastructure (manholes and milestones count, average altitude) and users. Each building – a sub-physical space with physical boundary – has equipment/tools and users, and is described by size, type, purpose and distance from the NOC. The size gives the total area of the building while type specifies the class of the building according to the number of storey (or floors). Purpose offers a description of the aim or major task performed in the building and may include any of the following: lecture (classroom, lecture theatre, auditorium), study (library), practical demonstration/experiment (laboratory), administration (office), residential (hostel), social activity and recreation (restaurant, eatery, staff club) and commercial (business centre, multipurpose, cyber café, bookshop, retail shop), store (machine, animal, good/item). Equipment/tools may have any of the following types – electronic, mechanical, and wooden; and has the following categories: computer/peripheral (UPS, router, mouse, other related device), instrument (laboratory equipment) and others (television, radio, and other related device). A summary of the FO infrastructure in the five campuses is given in Table 3. Currently, only three out of the five campuses have operational WLAN infrastructure with NOC and BS infrastructure in each campus. The central hub is the NOC, and each building and is defined by the attribute set (type, size, purpose, distance).

Table 3. Summary of the FO infrastructure in the five campuses

<table>
<thead>
<tr>
<th>Campus</th>
<th>FOC</th>
<th>MH</th>
<th>OFC</th>
<th>ZMH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Average Altitude</td>
<td>Count</td>
<td>Average Altitude</td>
</tr>
<tr>
<td>TC</td>
<td>17</td>
<td>77.00</td>
<td>20</td>
<td>83.50</td>
</tr>
<tr>
<td>TCAX</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>MC</td>
<td>29</td>
<td>61.50</td>
<td>13</td>
<td>61.80</td>
</tr>
<tr>
<td>UUTH</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>BSC</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

4.3.2 Technological Infrastructure

Technological infrastructure component of the test bed comprises computer resources and network infrastructure (NI). The NI includes hardware, software, service, user and NOC components. Network services are expected products of the test bed and include the following categories: T-1 Line, DSL, satellite, wireless protocols and IP addressing. The Computer resources are used by end-users to harness the network infrastructure provided in test bed. These include hardware — ranging from microcomputers (standalone desktops, laptops, palmtops and peripherals) and other devices. Software resources are systems and applications software such as operating systems (OSs). NI holistically defines the entire computer resources (hardware, software and users) that provides and facilitates network connections, communications, operations, services and management of the entire network resources. It enables links and provides network services between users, applications, services and other network resources (including internet, intranet and extranets). Network hardware represents the physical resources of the network. It includes servers, computers, data centres, LAN cards, switches, hubs and routers, etc. Network software tools are necessary to coordinate, plan and manage the entire network resources. The major tools in our test bed include network operations and management (NOM), network operating systems (NOSs), firewall and network security applications (NSAs). These are basic tools used by the NOC and data centers in the coordination and management of the entire network resources. This class of software provides means of monitoring enterprise resource planning (ERP), customer relationship management, and productivity applications, among others. NOS features include protocol support, processor support, hardware detection and multiprocessing support for applications. In addition, NOSs pose security issues such as authentication, restrictions, authorizations and access control as well as file, web service, printing and replication. The NOS performs tasks involving users’ administration, system and file management, and maintenance activities like backup, audit trails, system diagnostic, folder creations etc. Further tasks include security monitoring on all resources in the network and setting of priority on other network and computer resources. The distribution of hardware and software devices in each of the physical space is given in Table 4.
Table 4. Distribution of hardware/software devices in the three major campuses

<table>
<thead>
<tr>
<th>Hardware/software</th>
<th>Main campus</th>
<th>Town campus</th>
<th>Town annex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backup satellite dish</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Radio (FO last mile bridge – point-to-point, Access Point)</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Media converter</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Router (Huawei, CISCO, DHCP, MikroTik)</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Switch (distribution, access, FO)</td>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Antenna (sectorial)</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Base station</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wireless AP</td>
<td>6</td>
<td>40</td>
<td>22</td>
</tr>
<tr>
<td>Patch panel</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Storage device</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Server (SQL, active directory, CISCO business 6000 edition)</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Firewall (Huawei)</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4.3.3 NOC Infrastructure

The NOC infrastructure is required to efficiently manage the available network resources for all clients/users and report on the performance using service quality measures. We document the expected NOC infrastructure as follows:

*Dude server:* provides server configuring capabilities and other features such as layout (image map) of the entire network, coordination, management and monitoring of activities and services on devices and execution of actions based on devices’ state modification.

*Campus site map server:* provides an interactive virtual network that allows users the ability to interact with the NOC network resources. This server is configured to manage the various physical spaces within the entire campus network.

*Gateway server:* is a rack mounted map layout employed for the provision of remote monitoring and configuring of all the main routers in the study area that obtain public IP addresses via its virtual environment within the various campuses.

*Network WiFi server:* is a rack mounted map layout that provides remote monitoring and configuration of all WiFi access points within buildings, across the various campuses (using its virtual environment). This server manages the WiFi network performance in the buildings and depends on the test bed network QoS values.

*Portal map server:* is a rack mounted map layout that provides remote monitoring and configuring of all connection and request made at the campus/university portal via the virtual environment.

*Access Points (APs) map:* is a map showing the wireless transmitting points within the physical spaces, irrespective of the attribute set.

4.4 RSSI Capture

To capture the received signal strength information of the service area, a scan of the study environment (where APs are located) was performed using the Acrylic WiFi Professional – a WiFi analyzer software that identifies access points and WiFi channels, and useful for analyzing and resolving incidences on 802.11a/b/g/n/ac wireless networks in real time.

The functionalities of Acrylic include the following,

(i) efficient visualization of wireless network performance and connected users;
(ii) access point data transmission speeds identification, and channels optimization;
(iii) access WiFi network detailed information collation and visualization, including hidden wireless networks.

A scan of the environment was delayed for about two to three minutes to allow for full devices detection. The detected infrastructure and measurements were finally exported to a comma separated value (CSV) file and
A naive design simulation of three APs with one NOC is shown in Figure 6. Here, the challenges inherent in the deployed network become more apparent. First, closer APs may overlap and present security vulnerabilities. Second, radio signal interference and multipath propagation between APs, mobile users (MUs) and fixed users (FUs) may occur as a result of signal transmission between the various devices. Third, devices may roam for services, and create interoperability issues – characteristic of a system which interfaces are completely known to work with other systems – present or future in either implementation or access and without any restrictions.
The proposed location-aware framework indeed shall offer autonomous calibration and re-configuration to deal with issues connected with radio signal interference and multipath propagation. As regards security vulnerabilities, each AP is accounted for by exploiting inherent properties of the wireless channel – where communication channels can be monitored in real-time, and channel accountability made transparent through efficient channel visualization. To arrest the problems of roaming and interoperability, each transmitter requires a unique signalprint to detect illegitimate spoofing of client’s MAC addresses and for effective localization of compatible communication interface.

5. Service Quality Modeling

Service Quality (SQ) compares the perceived expectations of a service with the perceived performance. To achieve the perceived (expected) SQ, field trials were performed at the three campuses and the requirements as well as challenges for developing suitable test bed infrastructure, appraised. These field trials and preliminary discoveries necessitated the development of a suitable model to drive and optimize the perceived performance of the network. The modeling process involved three major processes: feature abstraction, feature pattern analysis and signalprint representation. These processes are discussed in the following subsections.

5.1 Feature Abstraction

Principal Component Analysis (PCA) was used to abstract the relevant features. The central idea behind PCA is
to reduce the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much as possible, variation present in the dataset (Jolliffe, 1986). We achieved this by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered such that the first few components retain most of the variation present in all of the original variables. The steps taken to achieve the PCA are summarized in four steps as follows:

(i) **Dataset normalization:** all features must be represented on the same scale, by putting each feature on a normal curve – i.e., compute the z-score of each feature $z = \frac{x_i - \overline{x}}{\sigma_x}$, where $x_i$ is the feature set, $\overline{x}$ is the mean of the feature set, and $\sigma_x$ is the standard deviation;

(ii) **Compute the covariance matrix:** The covariance between two variables, X and Y, can be given by the formula:

$$cov = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{(n-1)}$$

where $n$ is the sample size. This matrix is symmetric ($A = A^T$), and has a diagonal of just variances, because $cov(x,x)$ is the same thing as the variance of $x$;

(iii) **Find the eigen vectors (a non-zero vector whose direction does not change when a linear transformation is applied to it) of the covariance matrix:** $Av = \lambda v$; where $\lambda$, is a scalar, known as eigen value, associated with the eigen value, $v$;

(iv) **Transform the data to be in terms of the components:** This was achieved by multiplying the transpose of the feature vector and the transpose of matrix containing the data.

5.2 Feature Pattern Analysis

The Self-organizing map (SOM) was used to model the extracted patterns (from PCA). SOM is a data visualization technique, which reduces the dimensions of data through the use of self-organizing neural networks. The problem that data visualization attempts to solve is that humans simply cannot visualize high dimensional data; hence, techniques are necessary to assist the understanding of this high dimensional data. SOMs accomplish two things: they reduce dimensions and display similarities. In this paper, SOM was adopted to observe the inherent patterns exhibited by the resulting dimension reduction from PCA. The basic SOM is as a nonlinear, ordered, smooth mapping of high-dimensional data manifolds onto the elements of a regular, low-dimensional array, and this mapping is implemented as follows:

(i) define the set of input variables $x_j$ as a real vector, $x_j = (x_1, x_2, ..., x_n)^T \in \mathbb{R}^n$;

(ii) associate with each element in the SOM array a parametric real vector: $m_i = (m_1, m_2, ..., m_n)^T \in \mathbb{R}^n$ – also called a model;

(iii) define a distance measure between $x$ and $m_i$, denoted as $d(x, m_i)$.

The incremental SOM algorithm with a rectangular map topology, as shown in Figure 7, is adopted for the pattern analysis.

The image of an input vector $x$ on the SOM array is defined as the array element $m_e$ that best matches with $x$, with the index $\arg \min_i (d(x, m_i))$. The main task here is to define $m_i$ such that the mapping is ordered and descriptive of the distribution of $x$. The process in which such mappings are formed is defined by the SOM algorithm. This process is then likely to produce asymptotically converged values for the models, $m_i$, a
collection of which will approximate the distribution of the input samples $x(t)$, even in an ordered fashion. Vector $x$ may be compared with all the $m_i$ in any metric. In many practical applications, the smallest of the Euclidian distance $|x - m_i|$ defined by the subscript, can be used to define the best matching node (BMN):

$$c = \arg \min_i ||x - m_i|| = ||x - m_c|| = \min_i.$$  (3)

5.3 Signalprint Representation

In this study, the IT2FL (Mendel and John, 2002; Mendel and Liu, 2007) system was exploited to create signalprints for each infrastructure (be it fixed or mobile). The goal is to efficiently model uncertainties inherent in the field data – for accurate estimation and production of acceptable service quality with minimum uncertainty. To eliminate the drawbacks of any individual variable, fuzzifiable RSSI parameters were characterized. The proposed IT2FL-WLAN-SQ framework is shown in Figure 8, and comprises of five major components namely; fuzzifier, knowledge base, inference engine, type-reducer and defuzzifier.

5.4 Service Quality Optimization

The Adaptive Neuro Fuzzy Inference System (ANFIS) – a graphical network representation of Sugeno-type fuzzy system endowed with neural learning capabilities was used in this study for the purpose of learning the captured dataset, and predicting with high precision the service quality experienced at the various locations within the service area. ANFIS provides the method for the fuzzy modeling procedure to learn information about the RSSI and building dataset, and computes the membership function parameters that optimally associates the fuzzy inference system to signal strength changes that occur at the various APs. The system prediction gathers information from the knowledge-base, and is driven by the fuzzy “if then” rules (Takagi and Sugeno, 1983) to obtain accurate predictions.

Five common layers form the ANFIS architecture. The layers descriptions are as follows:

Layer 1 (L1) – each node generates the membership grades of a linguistic label. For the purpose of this study, the input nodes used are as defined in Table 5. We started with a larger set of features and abstracted away less significant features to generate more compact and robust feature set that will ensure accurate prediction.

Layer 2 (L2) – each node computes the firing strength of each rule using the min or prod operator. In general, any other fuzzy AND operation is used.

Layer 3 (L3) – nodes calculate the ratios of the rule’s firing strength to the sum of all the rules firing strength. The result is a normalized firing strength.

In Layer 4 (L4) – nodes compute a parameter function on the output of layer 3. Parameters in this layer are called consequent parameters.

Layer 5 (L5) – usually a single node that aggregates the overall output as the summation of all incoming signals. Figure 9 Illustrates the architecture for a $p$ input ANFIS.
6. Results

6.1 PCA Dimension Reduction

Table 6 shows Eigen values of the input parameters (RSSI and site-specific parameters) with their respective contributions to the perceived service quality. Components with Eigen value of at least one (1) may be accepted as principal component (Zou, Hastie and Tibshirani, 2006). As can be seen in the table, out of the nineteen parameters, six of them gave Eigen values above one while two parameters had Eigen values approximately one. Hence eight parameters were selected, and these parameters include: building height (BHeight), building type (BType), numbers of rooms (NOR), building size (BSize), number of floor(s) (Floor), building purpose (BPurpose), pathloss (Pathloss), and distance from network operation centre (DFNOC). Hence, contributing to 79.59% of the cumulative Eigen values (total parameters considered).

Table 6. Eigen values of input parameters with their cumulative percentage

<table>
<thead>
<tr>
<th>PC</th>
<th>Parameter</th>
<th>Eigen value</th>
<th>Difference</th>
<th>Proportion (%)</th>
<th>Histogram</th>
<th>Cumulative (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BHeight</td>
<td>3.721831</td>
<td>1.011952</td>
<td>20.68 %</td>
<td></td>
<td>20.68 %</td>
</tr>
<tr>
<td>2</td>
<td>BType</td>
<td>2.709880</td>
<td>0.253909</td>
<td>15.05 %</td>
<td></td>
<td>35.73 %</td>
</tr>
<tr>
<td>3</td>
<td>NOR</td>
<td>2.455971</td>
<td>1.117930</td>
<td>13.64 %</td>
<td></td>
<td>49.38 %</td>
</tr>
</tbody>
</table>
A scree plot explaining the fraction of the total Eigen values contributed by all the input parameters is shown in Figure 10. The contributions are arranged in descending order of cumulative contributions, and along the red line, a calibration of the numerical value of each principal component (PC) is achieved. The ‘elbow’ is at PC4, indicating that PC1, PC2 and PC3 are the most prominent parameters.
6.2 SOM Pattern Visualization

The U-Matrix clustering map (with service quality labels) representing captured data from the three major campuses under study is presented in Figure 11. To satisfy a preliminary field trial assessment and uncover the overall (perceived) performance, the service quality was grouped into three clusters (P-Poor, G-Good and E-Excellent) for the eight parameters selected through PCA. Distance of the adjacent neurons were computed and shaded as different colors between the adjacent nodes. Whereas dark colors between the neurons correspond to a large distance and are regarded as boundary separators, light colors between the neurons indicated close distance and are regarded as clusters. From the map, we observed that the overall perceived performance was fair, as good signals (G and E) dominated the map.

![Figure 11. U-Matrix Clustering Map](image)

Figure 12 is a combined visualization of the U-Matrix and feature component planes selected. This structure assists in explaining the clustering structure and similarities between the variables (Lin, Brusilovsky and He, 2011). A darker shade corresponds to larger ordinal values, a gray shade represents medium ordinal values, and a lighter shade represents low ordinal values of operation sequence. We observed that the SOM classified the feature component planes into four similarity groups according to the degree of clustering achieved. *Similarity group 1*: (building height (B_He) and building type (B_Ty), showed strong correlation with well clustered patterns, and were the first two highest PC values obtained.*Similarity group 2*: number of rooms (NOR), building size (B_Si) and number of floor(s) (N_Fl), also showed strong correlation with moderate cluster patterns, but the clusters of NOR and B_Si correlated better than N_Fl.*Similarity group 3*: building purpose (B_Pu) showed well separated clusters, but stands out on its own. *Similarity group 4*: Pathloss (P_loss) and distance from NOC (DFNO), showed very weak correlation. Hence, the SOM results validate the authenticity of our PCA algorithm, and proves that indeed pathloss and DFNO contributed less to the service quality.
6.3 ANFIS Service Quality Optimization

ANFIS consists of two major components: a fuzzy inference system (FIS) and an ensemble of artificial neural networks (NNs). Since it integrates both neural networks and fuzzy logic principles, it has the potential to capture the benefits of both in a single framework. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions. Hence, ANFIS is considered to be a universal estimator (Jang, Sun and Mizutani, 1997). Given input and output datasets, ANFIS constructs fuzzy membership function parameters (adjusted by a hybrid learning algorithm) to approximate precisely, the model parameters. The hybrid algorithm of ANFIS combines the gradient descent and least square methods. The supervised learning stage was actualized by selecting fuzzifiable input parameters, and five fuzzifiable parameters were identified (BHeight, BSize, NOR, Pathloss and DFNOC) and used to build the fuzzy rules. The generated signalprints (SQ) from the IT2FL system were used as target parameters. For efficient SQ optimization, the input parameters were further extended to five linguistic terms or membership functions (Excellent, Very Good, Good, Poor and Very Poor), as the initial SQ clustering was not optimal (see Figure 11). The membership functions are necessary to cause interaction of the input parameters during the training phase, until an optimal performance is achieved. The membership functions for the input parameters as generated by the ANFIS are presented in Figure 13.
The Sugeno fuzzy inference system (useful for modeling nonlinear systems by interpolating between multiple linear models) was adopted, and the dataset were partitioned into: 70% training sample, and 30% testing sample, and distributed according to samples× features (S×F) as shown in Table 7.

Table 7 Distribution of dataset for training and testing

<table>
<thead>
<tr>
<th>Category</th>
<th>Town Campus</th>
<th>Annex Campus</th>
<th>Main Campus</th>
<th>Partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs</td>
<td>2067×5</td>
<td>925×5</td>
<td>291×5</td>
<td>100</td>
</tr>
<tr>
<td>Targets</td>
<td>2067×1</td>
<td>925×1</td>
<td>291×1</td>
<td>100</td>
</tr>
<tr>
<td>Train inputs</td>
<td>1447×5</td>
<td>648×5</td>
<td>304×5</td>
<td>70</td>
</tr>
<tr>
<td>Train targets</td>
<td>1447×1</td>
<td>648×1</td>
<td>304×1</td>
<td>70</td>
</tr>
<tr>
<td>Test inputs</td>
<td>620×5</td>
<td>278×5</td>
<td>87×5</td>
<td>30</td>
</tr>
<tr>
<td>Test targets</td>
<td>620×1</td>
<td>278×1</td>
<td>87×1</td>
<td>30</td>
</tr>
</tbody>
</table>

6.4 Performance Evaluation

The performance evaluation metrics used to evaluate the system performance are the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). The RMSE measures the differences between the (sample and population) values predicted by a model or an estimator and the values actually observed, and is given by the expression,

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{expected} - x_{observed})^2} = \sqrt{\text{mean}(e_i^2)}$$  \hspace{1cm} (2)

Where \( x_{expected} \) are the service quality values in the dataset, and \( x_{observed} \) are the resultant target output. The MAE measures the accuracy of a forecast or model (Hyndman and Athanasopoulos, 2014), and is given by the expression,

$$MAE = \frac{\sum_{i=1}^{n} |e_i|}{n} = \text{mean}(|e_i|)$$  \hspace{1cm} (3)

The error signals use a back-propagation algorithm while tuning the premise parameters. The gradient descent algorithm is implemented during the backward pass phase. Whereas, only the consequent parameters changed during the forward pass, they remained unchanged during the backward pass. The iterations of the forward and backward passes caused the premise and consequent parameters to be identified and generated for the FIS. Figures 14, 15, 16 and 17, present the error plots and relationship between target and predicted output for both trained (a) and test (b) dataset in the town campus, town campus annex, main campus, and the three campuses.

Figure 13. ANFIS membership function segmentation
combined, respectively.

Figure 14. Train and test data for town campus

(a) Train data  
(b) Test data

Figure 15. Train and test data for town annex campus

(a) Train data  
(b) Test data

Figure 16. Train and test data for main campus

(a) Train data  
(b) Test data
The results revealed that the ANFIS model was sufficient for optimizing the SQ, as the RMSE values of training and test data gave close results for all the study sites considered (i.e., town campus (19.81 and 19.52), town campus annex (21.30 and 22.24), main campus (16.95 and 16.58) and all campuses combined (20.11 and 20.42)). Also, training result for all campuses combined produced the least MAE of 0.0979, but with the largest MAE for testing (0.7345). However, the least MAE value came from town campus Annex (0.4707). The wide MAE deviations between the training and test data may not be unconnected with the high degree of uncertainties experienced by the system under study, as well as the quality of data collected. Interference between mobile and fixed infrastructure may have also posed challenges, hence leading to the wrong classification of (some) features.

7. Conclusion, Recommendation and Future Research Direction

WLAN technology has advanced to become a practical alternative to traditional networks, and offers seamless connectivity through well installed APs. A successful WLAN implementation implies striking the right balance between functionality, performance and security objectives, and can best be achieved through careful network planning and interaction with experts. In this paper, intelligent tools were explored to abstract features that show sufficient contribution to the perceived service quality, required to optimize the system performance.

The following recommendations are proffered as measures for minimizing poor service quality in academic environments:

(i) **Detailed site planning and AP relocation**: Wide difference in ME (between trained and test data) suggests poorly deployed WLAN infrastructure. Hence, a detailed site planning with specific attention to signal transmission between the respective infrastructure devices is necessary. This will minimize multi-path propagation (an important contributor) which was judged as being less important, by both PCA and SOM;

(ii) **Resource aggregation**: Uniting the various resources at different campuses can improve the service quality and save cost. The LAS feature taxonomy designed in section 4.2, could serve as a useful metadata for a real life system deployment;

(iv) **Efficient mapping of perceived service quality performance with expected performance**: This process is important to re-evaluate the overall service quality, after achieving the first two tasks.
A lot can be added to the location-aware framework in the future, as the framework can be extended to provide real time location based service through the introduction of wireless sensor nodes (WSNs) – at each of the access points – for monitoring and live updates of ‘mission critical’ services.

Acknowledgments

We appreciate the support of the Tertiary Education Trust Fund (TETFund) – for funding the research; our Undergraduate and Postgraduate students and the ICT personnel of the University of Uyo – for their involvement in the fieldwork phase of this research (test bed construction, field measurements, and deployment).

References


that Support the User. *OECD Publishing, France.*


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).