**IOT Based Real-Time Customer Retention Tracking and Comprehensive**

**Analysis using LangChain & LangGraph**

Prof. Sheela Chinchmalatpure, Anurag Raut, Ridham Anand, Rimaz Bhaldar, Rohan Khanna

**Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India**

[[1]](#footnote-2)

***Abstract*** *—* Customer loyalty is important for the continued existence of business in general and, particularly for large exhibitions and events, because of the difficulty involved in ensuring the attendees visit all the stands. On the other hand, such traditional methods as research and observation analysis are usually insufficient, prone to errors, and devoid of implementing insights in practice. The project proposes a solution in the form of an IoT-based real-time customer binding system and an advanced analytical oversight based on Langchain and Langgraph for intelligent data processing and decision-making. Such insights are processed in cloud-based analytics so that the companies can analyze engagement metrics and effectively optimize stand layout. Chatbots built with the help of Latchain and Langgraph, equipped with an AI mechanism, allow interesting groups to easily inquire about the data and get implementation able recommendations without requiring in-depth technical knowledge. Built to be scalable and easy to deliver, it complies with industry demands for AI-operated optimization and customer experience.

***Key words:*** Customer Retention, Real-Time Tracking, Nordic BLE 4.0, ESP32, LangChain, LangGraph, Cloud Analytics, AI Chatbot, Wearable Devices, Booth Sensors, Data-Driven Insights, Visitor Engagement, AI-Powered Analysis

# ***Introduction***

Customer loyalty plays an important role in long-lasting success in trade fairs and events in today's competitive business environment. Hence, knowledge about the pattern of engagement of the visitors becomes a part of the optimization process with respect to stand layout, product placement, and marketing strategies. However, traditional techniques, be it manual inquiry, casual investigation, RFID detection, or thermal imaging, rarely accomplish real-time knowledge or reasonable data in general. The project proposes an IoT-based loyalty and analytics system with real-time customer constraints and advanced analytics. The technology monitors visitor interactivity and engagement within stands using an ESP32 microcontroller with a Nordic BLE 4.0 chip integrated to track movement patterns through portable devices and, eventually, commodity stands for seamless data collection and transmission. Such real-time knowledge processing loads on to cloud-based analytics to provide the company with a more in-depth understanding of consumer behaviour. Built with LangChain and LangGraph, AI enabled Chat Bot allows the dynamic query of data and easy access to insights without any specialized technical knowledge. The solution proposed is scalable and efficient, and developed in accordance with the modern-day customer loyalty strategies.

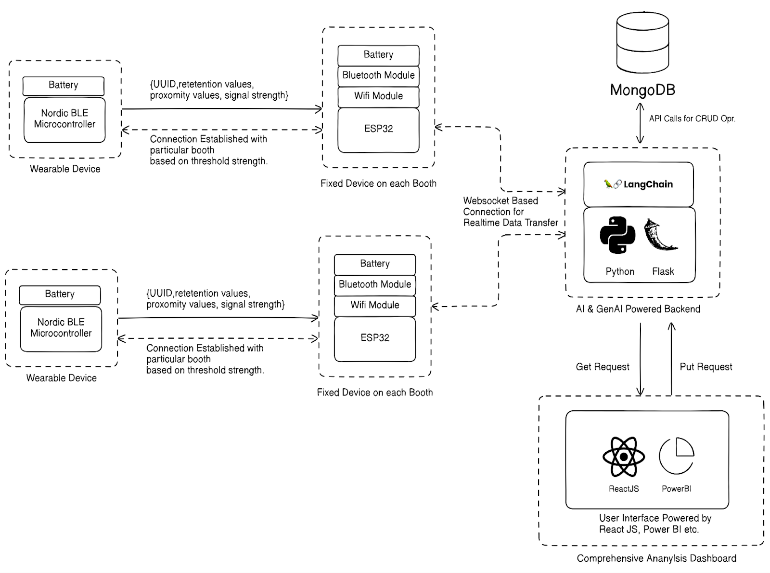
# ***Methodology***

The proposed system for tracking and analyzing customer retention integrates a Bluetooth Low Energy (BLE) wearable device, an ESP32-based edge computing unit, and a real-time data transmission pipeline to monitor visitor movement efficiently.

**A. System Components and Architecture**

**1. Proposed System Architecture**

The system architecture comprises **wearable devices, fixed booth devices, a backend powered by AI, and a real-time analytics dashboard**. It enables **crowd movement tracking, engagement analysis, and intelligent insights** for exhibitions.

****

*Fig.1 Proposed System Architecture*

In Fig.1: The proposed system architecture consists of four primary layers:

1. Wearable Device Layer – Nordic BLE microcontroller-based devices transmit UUIDs, retention values, proximity data, and signal strength to detect visitor interactions with booths.
2. Fixed Booth Infrastructure – Each booth is equipped with ESP32 devices (Wi-Fi and Bluetooth modules) that receive wearable device signals and transmit data via a WebSocket-based real-time communication system.
3. AI-Powered Backend – A Flask-based backend integrates LangChain for AI-driven analytics, with MongoDB handling CRUD operations for visitor engagement tracking.
4. Frontend & Analytics Dashboard – The UI, built with React.js and Power BI, provides real-time visual analytics and supports a business executive-focused AI chatbot for insights.

This scalable, real-time framework facilitates intelligent visitor tracking and behavioral analysis in exhibitions.

**2. Wearable BLE Device (Nordic Microcontroller)**

Each visitor is equipped with a BLE-enabled wearable, developed using a Nordic nRF52840 microcontroller, which continuously advertises its presence to nearby ESP32-based receivers.

**2.1 BLE Advertisement Parameters**

* TransmissionPower**:** =

(set using Bluefruit.setTxPower(4))

* AdvertisingInterval**:**

ms, (converted from (32, 244) \* 0.625 ms

* Data Format**:** Includes a UUID, device ID, and a counter.

The device continuously transmits advertising packets containing a custom BLE service UUID and a characteristic value. The transmit power is set at +4 dBm to maintain signal consistency.

**3. ESP32-Based Booth Node**

Each booth is equipped with an ESP32 module configured to scan for BLE packets, compute real-time RSSI-based distance estimates, and log connection duration. The mathematical modelling of RSSI-to-distance conversion and signal processing ensures robust localization.

**3.1 BLE Scanning & Signal Processing**

* ESP32 operates in active scanning mode, capturing BLE advertisements at 1-second intervals.
* Only packets with the expected UUID signature are processed, filtering out noise from non-system devices.
* The RSSI (​) values are extracted and recorded alongside timestamps.

**3.2 RSSI-Based Distance Estimation & Thresholding**

The system estimates the distance between the wearable and an ESP32 booth node using the log-distance path loss model (LDLPM):

where:

* .
* .
* .

To reduce RSSI fluctuations due to interference, the system applies an Exponential Moving Average (EMA) filter:

where:

**3.3 Disconnection Handling**

If RSSI remains below −70 dBm for multiple consecutive readings, the ESP32**:**

* Disconnects the wearable device.
* Transmits the readings to the backend.

The processed data is sent to the backend **via HTTP POST in JSON format**, ensuring smooth integration with analytics and tracking systems.

HTTP POST Request contains the following:

* Booth ID
* In and Out time of Wearable Device
* RSSI Values
* Distance Values
* Wearable Device ID (UUID)

**B. Data Processing and AI Analysis**

**1. Backend System Methodology**

The backend system processes visitor movement data to compute dwell times, classify visitor engagement levels, analyse foot traffic patterns, and generate engagement scores. It is implemented using Flask as the web framework, MongoDB for NoSQL storage, and Flask-SocketIO for real-time updates. The system ensures efficient data ingestion, aggregation, and retrieval using indexed queries, event-driven processing, and structured mathematical modelling.

* 1. **Mathematical Modelling & Data Processing**

1. **Dwell Time Computation:**

Dwell time is the duration a visitor remains within a booth or zone:

where timestamps are converted to minutes for standardization.

Visitors are segmented based on dwell time:

* Brief Visitors: Dwell Time <5 min
* Average Visitors: 5≤ Dwell Time<15 min
* Engaged Visitors: Dwell Time≥15 min

These classifications are stored in MongoDB for analytical retrieval.

1. **Visitor Segmentation by Dwell Time**

For each segment, visitor counts are derived using an **indicator function**:

* )
* )

1. **Hourly Visitor Distribution**

Visitor counts are aggregated per hour to identify peak traffic times:

for h∈[0,23], mapping each visit timestamp to its respective hour bucket.

Stored in MongoDB (trafficByHour) for time-series analysis.

1. **Dwell Time Distribution (Binning Method)**

A **5-minute binning approach** is used for dwell time histograms:

where represents the bin index.

Data is stored in dwellTimeDistribution, aiding engagement trend analysis.

1. **Engagement Score Computation**

Engagement scores are derived from multiple behavioural metrics:

Weights are empirically determined based on business needs.

The engagement scores are stored in MongoDB (engagementScores) for comparative analysis.

1. **Return Rate Calculation**

Return rate determines how many visitors came back on different days. Given (repeat visitors) and (total visitors):

**2. AI analysis using LangChain & LangGraph**

The system leverages LangChain and LangGraph to enable multi-turn AI-driven analytics for business executives. The chatbot provides data-backed insights using real-time and historical visitor analytics, stored in a vector database (Pinecone). It employs Google Vertex AI (gemini-1.5-flash) as the LLM backbone for generating responses and performing semantic search over structured and unstructured data.

**2.1 Embedding Representation and Vector Storage**

The system embeds structured and unstructured attendee data (visit durations, booth interactions, historical trends) into 768-dimensional vectors using Google Vertex AI's text embedding model:

where is the embedding function, and represents a document containing attendee insights. These embeddings are stored in Pinecone, enabling semantic search over retention trends.

**2.2 Vector Search for Retrieval-Based Analytics**

The chatbot utilizes vector similarity search to fetch relevant attendee retention patterns and business insights:

where:



By using cosine similarity, the system accurately matches executive queries with real visitor behaviour data, enhancing decision-making efficiency for business executives.

**2.3 LangGraph for Structured Workflow Execution**

The retrieval and response generation follows a graph-based processing model:

where:

* e representing key workflow steps (retrieval, trend analysis, dwell time computation).
* are are edges defining conditional execution paths based on query intent.

This enables scalable and modular execution of analytics requests, ensuring real-time responses.

**2.4 Retrieval-Augmented Generation (RAG) for Business Insights Generation**

The chatbot employs RAG-based response generation, where retrieved analytics data is fused with Google Vertex AI-generated insights:

where:

* represents the probability of generating a response based on historical booth interactions and visitor trend.
* represents the probability from Google Vertex AI's model.
* is a hyperparameter controlling the weighting of factual retrieval vs. AI-generated insights.

This method ensures accurate, context-aware, and insightful responses, making the chatbot highly effective for business executives.

**C. Data Visualization and Knowledge Generation**

The frontend is developed using React.js, ensuring a modular and responsive user interface. It integrates with the backend via RESTful APIs and WebSockets to enable real-time updates. The system is structured into multiple components, each serving a specific function.

1. **Real-Time Data Visualization**

* The frontend fetches visitor data via RESTful API calls to the Flask backend and renders it dynamically using Recharts.
* Dwell time, return rates, traffic trends, and engagement scores are computed and displayed using bar charts, pie charts, and line graphs.
* Booth-wise performance metrics (footfall, average dwell time, engagement scores) are analyzed for comparative insights.

1. **Customizable Metrics Selection & Data Sorting**

* Users can sort data based on visitor count, average dwell time, engagement score, and return rate.
* Sorting and filtering operations are handled efficiently using React state management (useState, useEffect).

1. **Business Executive Chatbot**:

* Uses LangChain and Google Vertex AI for conversational AI.
* Provides actionable insights, answering questions related to visitor trends, booth performance, and recommendations.
* Accepts natural language queries for real-time analytics interpretation

1. **Real-Time Data Updates**

* Implements Socket.IO for instant synchronization of visitor analytics.
* Reflects changes in data (e.g., new visitor counts, updated engagement scores) without requiring page reloads.

# ***Literature Review***

Customer loyalty was examined in detail in marketing, HR and business analysis. The AI control predictive analysis and commitment strategies for RFV retention have been studied for many years, among them Madanchian (2024), who discusses AI tools for decision-making in HR and Kemparaju et al. (2023) investigating applications for machine learning for employee loyalty. Basnet (2020) examines the role played by AI in predictive analytics for customer loyalty. Ascarza et al. (2019) address data-controlled retention management analysis, while Letting (2018) exhibits CRM practices. The study expands on all of the above by building on them with IoT-based real-time tracking and AI analytics as an improvement in customer loyalty. Using Langchain and Langgraph, dynamic data processing and interactive query technology allows traditional retention analytics to handle gaps.

# ***Results and Discussion***

The evaluation of the customer retention tracking system for exhibitions highlights its effectiveness in real-time visitor tracking, behavio ural analysis, and executive decision support. The system provides data-driven insights that help exhibitors understand visitor engagement patterns, optimize booth strategies, and enhance business decision-making.

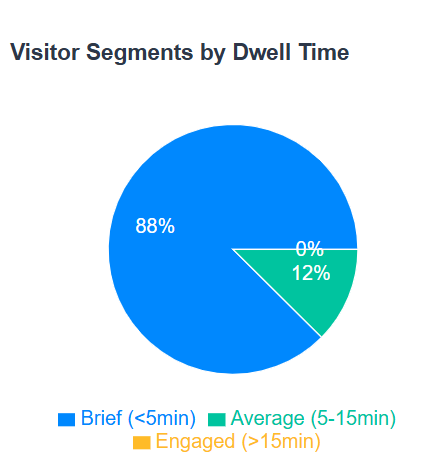
1. **Visitor Distribution and Engagement Segmentation**

Visitors are classified into three categories based on their dwell time:

* Brief Visitors (<5 min): Passersby with minimal engagement.
* Average Visitors (5–15 min): Moderate engagement, indicating interest.
* Engaged Visitors (>15 min): High-value visitors with a greater likelihood of conversion.

The distribution analysis shows that most visitors engage for short durations, while a smaller subset exhibits prolonged interest, which is critical for targeted engagement strategies.

(See Fig. 2: Visitor dwell time distribution and segmentation)



A graph with a bar and a number of bars

AI-generated content may be incorrect.

*Fig. 2 Dwell Time Distribution and*

*Segmentation*

1. **Peak Visitor Analysis and Temporal Trends**

Time-series analysis of visitor traffic reveals high engagement periods, allowing exhibitors to optimize staffing and promotional efforts. Key observations include:

* Peak traffic hours (12:00–15:00): Suggesting the best period for business interactions.
* Gradual decline post-evening: Indicating reduced engagement opportunities.

By leveraging these insights, exhibitors can adjust resource allocation and marketing strategies to maximize engagement during high-traffic periods.

*(See Fig. 3: Visitor footfall trends across time intervals)*

A graph with a line and numbers

AI-generated content may be incorrect.

*Fig. 3 Hourly Visitor Distribution*

1. **AI-Powered Executive Insights via Chatbot**

A LangChain and LangGraph-powered chatbot, integrated with Google Vertex AI, provides business executives with instant data-driven insights. Key functionalities include:

* Trend Analysis: Retrieves visitor trends and behavioural analytics.
* Comparison Insights: Benchmarks booths based on engagement performance.
* Real-time Queries: Enables executives to make informed decisions on visitor interaction

The chatbot significantly reduces manual effort in interpreting analytics, enabling faster and more efficient decision-making.

# ***Future Scope***

The proposed system also has a significant improvement and expansion potential. Future developments would need to see deep-learning AI models for more precise storage predictions. The system can be adapted for retail, hospitality, healthcare, and many other applications that enhance customer loyalty. Integration of Blockchain ensures secure data collection and data that is impossible to manipulate. Edge computing allows latency to be decreased by processing the data on-site with greater detail. Further additional explicit biometrics or surrounding sensors can capture slight nuances of customer behavior. Automated triggers for actions enable an immediate adjustment in stand layouts and marketing strategies based on real-time analysis which, in turn, enables fast adaptability of your system.

# ***Conclusion***

The project presented here develops a BLE-based real-time customer-binding tracking system integrated with AI-processed insights. The combination of portable BLE devices that are recipients of ESP32 on booth-based AI allows businesses to monitor visitor commitments at exhibitions and large events. Actual tracking and analysis provide a data control approach to optimize visitor experience and improve retention. The modular architecture ensures scalability and enables seamless integration across different industries. Future improvements such as deep learning for more precise behavioral analysis and expansion to other areas like retail and hospitality could further increase applicability. Incorporating better AI models and using blockchains could enhance the precision and reliability of the system for safe data processing. By overcoming current limitations present in traditional retention analytics, the solution will pave the way for a more efficient AI-controlled customer commitment and engagement.

***V. Acknowledgment***

We express our gratitude to Professor Sheela Chinchmalatpure for her valuable guidance and insights throughout the project. We would also like to thank our colleagues and faculty at the Vishwakarma Institute of Technology, Pune, for their guidance and constructive feedback. Special thanks go to the developers of Langchain and Langgraph, as this enabled the implementation of the AI control analysis systems.

***VI. References***

1. Mitra Madanchian, “From Recruitment to Retention: AI Tools for Human Resource Decision-Making (2024)” - Applied Sciences.
2. Prof. Kemparaju N, Ms. Latha JA, Ms. Likitha M B, Ms. Prerana Hiremath, Ms. Vaishnavi Ram J “Employee Retention Analysis Using Machine Learning (2023)” - International Journal for Research in Engineering Application & Management (IJREAM).
3. Sunil Basnet, “The Impact of AI-Driven Predictive Analytics on Employee Retention Strategies” - International Journal of Trend in Scientific Research and Development (2020).
4. Eva Ascarza, Scott A. Neslin, Oded Netzer, Zachery Anderson, Peter S. Fader, Sunil Gupta, Bruce Hardie, Aurélie Lemmens, “In Pursuit of Enhanced Customer Retention Management (2019)” - Journal of Service Research.
5. A. Letting, “The Effect of customer relationship management practices on customer retention in the exhibition industry within Nairobi City County (2018)” - International Journal of Scientific and Research Publications.
6. Bhenu Artha, Inna Zahara, Bahri, Niken Permata Sari, "Customer Retention: A Literature Review (2014)" - Journal of Business & Economics Research.
7. “Supporting Customer Retention through Real-Time Monitoring of Customer Complaints (2017)" - International Journal of Operations & Production Management.
8. Neeraj Kripalani, "Predictive Analytics for Customer Retention: A Data-Driven Framework for Proactive Engagement and Satisfaction Management (2021)" - Journal of Marketing Analytics.

1. [↑](#footnote-ref-2)