EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Development of a tourist behavior analysis system using public Wi-Fi data to enhance smart tourism management: A case study of Sri Chiang Mai Smart City

Thalerngsak Wiangwiset[1)], Chayada Surawanitkun[1)], Wullapa Wongsinlatam[1)], Tawun Remsungnen[1)], Kanda Sorn-in[1)], Apirat Siritaratiwat[2)], Chavis Srichan[3)], Prachya Thepparat[4)], Weerasak Bunsuk[4)], Aekkaphan Kaewchan[4)], Wisut Supasai[5)] and Ariya Namvong*[1)]

[1)]Center of Multidisciplinary Innovation Network Talent (MINT Center), Faculty of Interdisciplinary Studies, Department of Technology and Engineering, Khon Kaen University, Nong Khai Campus, Nong Khai 43000, Thailand
[2)]Faculty of Engineering, Department of Electrical Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[3)]Faculty of Engineering, Department of Computer Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[4)]Central Maintenance Sector, National Telecom Public Company Limited, Khon Kaen 40000, Thailand
[5)]Faculty of Science, Computer Center, Ramkhamheang University, Bangkok 10240, Thailand

Received 27 August 2024
Revised 17 December 2024
Accepted 25 December 2024

Abstract

Sri Chiang Mai Smart City faces challenges in tourism management due to its limited understanding of tourist behavior patterns and inefficient resource management at tourist attractions. The study aims to analyze tourist movement patterns and preferences across key locations, develop a privacy-preserving data collection framework, and provide data-driven recommendations for tourism resource allocation. The system analyzes Wi-Fi data from 11 strategic locations along the Mekong River using spatial and temporal analysis techniques. Tourist data is protected with SHA-256 encryption and tokenization of MAC addresses. Analysis of over 72 million connection records from May 2022 to December 2023 revealed significant patterns in tourist behavior. Tourist numbers showed dramatic increases during cultural events, with a 127.87% increase during the Songkran festival from April 13 to 15 compared to regular days in the same month at Boeng Wiang Courtyard. Four locations consistently showed higher visitor levels: Sri Chiang Mai Community Health Park, Naga Courtyard, In front of Had Pathum Temple, and Boeng Wiang Courtyard. Temporal analysis identified peak hours between 17:00 and 21:00, especially during the Songkran and longboat racing festivals. The findings enabled data-driven decisions in tourism resource allocation, event planning, and infrastructure development. However, the study is limited to tourists who connect to public Wi-Fi, potentially excluding those using mobile data or no digital devices. Despite these limitations, the system demonstrates how smart city infrastructure can effectively support smart tourism management through privacy-preserving data collection and analysis, contributing to the development of sustainable smart tourism initiatives.

Keywords: Smart city, Smart tourism, Tourist behavior, Wi-Fi usage analysis, Spatial and temporal analysis, Resource management

1. Introduction

Tourism plays a vital role in driving economic growth and development worldwide, especially in Thailand where it contributes significantly to the national GDP. The industry creates employment opportunities, stimulates local business growth, and promotes cultural exchange. In Thailand, tourism not only generates substantial revenue but also supports countless small businesses and local communities, making it a crucial sector for sustainable economic development.

In the era of smart city development, tourism management faces increasing demands for efficient and responsive service delivery. The widespread availability of free Wi-Fi in public spaces offers real-time insights into tourist behavior and movement patterns. This data-driven approach enables city planners to optimize infrastructure placement, allocate resources efficiently, and plan cultural events based on actual usage patterns. By integrating these insights into urban planning processes, cities can develop sustainable management strategies that balance tourist needs with local community interests.

Wi-Fi data analysis is crucial for efficient tourism management, economic growth, and community well-being in smart cities. Tourist attractions enhance destination vibrancy, support equitable development through accessibility, and provide environmental and cultural benefits. Well-managed areas strengthen local identity, boost businesses, and improve visitor experiences through data-driven insights [1-3].

Wi-Fi technology is crucial for analyzing tourist behavior through various innovative methods. Research has shown the effectiveness of Wi-Fi sensor devices in estimating real-time attendance at events with up to 95% accuracy by utilizing probe request messages and unique smartphone fingerprints [4]. Additionally, geolocation data from public Wi-Fi infrastructures has been successfully applied to monitor mobility and congestion in smart cities, aiding urban planning and management by estimating pedestrian density and social distancing levels in dynamic scenarios [5-7]. Privacy and ethical considerations are paramount when

*Corresponding author.
Email address: ariyna@kku.ac.th
doi: 10.14456/easr.2025.4

analyzing the utilization of Wi-Fi technology in public spaces due to risks associated with public Wi-Fi hotspots, such as privacy leakage, tracking behaviors, and data collection practices that could compromise sensitive user information [8-11]. While Wi-Fi tracking data can provide valuable insights for urban studies, it is crucial to consider its limitations and ethical implications, emphasizing the need to protect user privacy and ensure transparent data handling practices in public Wi-Fi environments [12-14].

The analysis of Wi-Fi data for tourism management raises several critical research questions:

- How can Wi-Fi usage data be effectively utilized to understand and predict tourist behavior patterns in smart city environments?
- What spatial and temporal patterns emerge in tourist activities across different locations and during various cultural events?
- How can privacy-preserving data collection methods be implemented while maintaining analytical value for tourism management?

This research has three primary objectives. The first is to conduct a comprehensive review of current research on tourist behavior analysis using public Wi-Fi data in a smart tourism context. The second objective is to serve as a data descriptor, detailing the methods for data collection, data processing, and data acquisition. This section will provide comprehensive guidelines on how to effectively gather, manage, and retrieve relevant data for analyzing tourist behavior through public Wi-Fi usage. The third objective is to utilize the dataset to analyze the patterns of tourist behavior and use the results of this analysis to offer recommendations for smart tourism management. By understanding how tourists utilize city spaces and facilities, the article aims to inform and guide tourism management decisions to enhance the efficiency of tourist services and experiences.

Building on previous work [15] that focused on the design and development of a crowd monitoring system, this article introduces an expanded dataset derived from our enhanced system. This growth was facilitated by upgraded access points that can support more users. We strategically relocated two installation sites to areas with higher traffic and extended our data collection phase from 5 to 14 months. With increased data, we introduced a redesigned database to optimize system performance. For data security, we anonymized and tokenized media access control addresses, hindering the potential for reverse tracking. The data and network security are managed by the National Telecom Public Company Limited (NT), ensuring robust protection. Accommodations were made in the system to manage the increased data volume. This dataset and associated source code have been made available on a public online platform to encourage wider use and research. As a result of these enhancements, our data assessment is more reliable, and the enlarged dataset has been used to fine-tune the evaluation process.

The remainder of the article is structured as follows: The Literature Review section covers existing research on tourist behavior analysis and Wi-Fi technology, identifying key findings and areas for further exploration. The Methods section explains the procedures for data collection, processing, and analysis, emphasizing privacy and ethical considerations in the research process. The Results section presents findings on tourist behavior. Finally, the Conclusion summarizes the findings, discusses their implications, offers recommendations for authorities, and addresses future research directions and study limitations.

2. Literature review

Understanding tourist behavior through Wi-Fi data analysis requires a comprehensive review of smart city concepts, analytical methods, and practical applications. This review explores the theoretical foundations of smart city development and spatial-temporal analysis to behavioral prediction, along with research evidence from Wi-Fi usage analysis, tourist forecasting, and resource management in various destinations. These interconnected aspects provide the framework for developing effective tourism management strategies.

2.1 Smart city theory

Analysis of smart city implementations reveals distinct approaches to Wi-Fi based tourist monitoring. The Yanfu Greenland Park study [16] represents the focused, single-location approach, successfully identifying specific usage patterns but was limited in scale. In contrast, municipal-level implementations like Turin [17] and the Madeira Islands [18] demonstrate broader coverage, with Madeira's network of 82 routers across 81 locations setting a benchmark for comprehensive monitoring. However, these large-scale implementations often sacrifice detailed behavioral insights for coverage.

Critical comparison of these approaches reveals an important trade-off: while single-location studies like Bologna [19] provide deep insights into specific urban interventions, they lack the broader context that city-wide implementations offer. This gap between depth and breadth of analysis remains a significant challenge in smart city research.

2.2 Spatial and temporal analysis theory

Analysis of spatial and temporal patterns in tourist behavior through Wi-Fi data reveals diverse approaches across international destinations. Research findings demonstrate a critical evolution from basic presence detection to sophisticated movement analysis.

Studies in Singapore's Marina Bay [20] represent an advanced event-focused analysis, where Wi-Fi probe requests during the i Light Singapore festival revealed not just crowd densities but specific behavioral adaptations to scheduled activities. This temporal-centric approach successfully identified how evening light shows influenced tourist distribution patterns, but lacked comprehensive spatial context.

In contrast, the Shichahai Scenic Area study in Beijing [21] achieved better integration of both dimensions. Their analysis distinguished different user groups' mobility patterns, revealing distinct preferences between holidays and weekends, while simultaneously mapping high-attraction zones near cultural sites. However, the study was limited by its inability to track complete movement patterns between zones.

The Oxford Circus research [22] advanced the field by introducing visibility analysis into Wi-Fi based tracking. Using nineteen nodes across an area with a 2-kilometer diameter, this study demonstrated how urban design elements influence pedestrian movement patterns. This integration of spatial visibility with movement data provided deeper insights into tourist decision-making, though it struggled with distinguishing between tourists and local traffic.

2.3 Behavioral prediction theory

Recent advances in tourist behavior prediction demonstrate a shift from traditional statistical methods to more sophisticated machine learning approaches, revealing both progress and persistent challenges in different analytical contexts.

The Kyoto study [23] pioneered an integrated approach in the Higashiyama Ward area, combining Wi-Fi sensor data with a recursive logit model to analyze tourist route choices. While this method effectively assessed time spent at locations and revealed patterns based on attraction types, it was limited by its reliance on historical data for predictions.

A significant advancement was achieved by Fantozzi et al. [24], who developed an innovative algorithm using sentence transformer neural networks to predict tourist visit intentions. Their approach achieved 90% accuracy in distinguishing visit intentions from social media posts, demonstrating robust performance even with imbalanced data. However, this method was dependent on active social media participation, potentially missing significant tourist segments.

Shrestha et al. [25] took a different approach by developing a personalized tourist recommender system. Their data-driven machine learning method analyzed multiple factors including demographics, behaviors, preferences, and satisfaction levels. While this comprehensive approach improved recommendation accuracy, it required extensive historical data and user input.

An Italian study [26] marked another significant advance, utilizing anonymized mobile data from Vodafone Italia with long short-term memory (LSTM) models. This approach proved more effective than traditional Markov models for predicting short-term foreign tourist movements. However, its dependence on mobile carrier data limited its applicability in areas with multiple service providers.

2.4 Research on analyzing Wi-Fi usage in tourist cities

Analysis of Wi-Fi usage research across tourist cities reveals evolving methodologies and varying levels of success in capturing and interpreting tourist behavior. Critical comparison of major studies highlights both technological advances and persistent challenges.

The Shichahai Scenic Area study [27] demonstrated sophisticated data collection capabilities, analyzing 670,000 tourist Wi-Fi probe requests. Their findings revealed clear correlations between location attractiveness and tourist dwell time, particularly near historical sites. However, while successfully identifying afternoon peak times and stay durations, the study struggled with distinguishing between repeat visitors and new tourists.

Research in Mallorca [28] advanced the field by developing more nuanced congestion analysis methods. This study revealed complex spatiotemporal patterns in pedestrian congestion, with significant variations between areas like Plaça Major and Passeig del Born during peak seasons. This work provided valuable insights into crowd dynamics but was limited by its focus on general pedestrian flow rather than specific tourist behavior.

A Lisbon study [29] took an innovative approach by differentiating between stable and volatile Wi-Fi networks for location verification. While achieving impressive 7-minute tracking accuracy and 89% match rates for indoor locations, the system's outdoor tracking capability dropped to 14%, highlighting the challenges of consistent outdoor Wi-Fi monitoring.

Florence's comprehensive study [30] represents one of the most extensive implementations, utilizing 345 carefully selected access points (APs) from a network of 1,500 APs. Their analysis of 56 million connection events over six months revealed that 60% of users were short-term visitors, tracking approximately 16% of the city's 14 million annual visitors. This research successfully identified about 34,000 daily users, with 10% being newcomers, though it faced challenges in distinguishing between tourists and local short-term visitors.

2.5 Research on forecasting tourist behavior

Recent advances in tourist behavior forecasting demonstrate a significant shift from traditional single-source analysis to more sophisticated multi-data approaches, revealing both opportunities and challenges in prediction accuracy.

A comprehensive review of tourism forecasting research [31] spanning the years 2012-2019 revealed critical insights into data source effectiveness. The study found that while Google Trends data dominated forecasting approaches, emerging AI techniques showed promising improvements in prediction accuracy. However, the research highlighted a significant gap between single-source and multi-source prediction models, suggesting that data integration remains a key challenge.

The Chinese study at Mount Siguniang [32] provided compelling evidence for device-specific analysis approaches. Their application of LSTM deep learning models revealed that mobile search queries outperformed PC searches in forecasting daily tourist arrivals, demonstrating an 8.55% improvement in accuracy. This finding highlighted the growing importance of mobile data in tourism forecasting, though the study was limited by its focus on a single search engine platform.

Research in Kerala [33] demonstrated the value of integrated data analysis, combining Google Trends data from 63 tourism-related keywords with traditional tourist statistics over a decade (2004-2014). Their identification of four key factors - attitude, perceived behavioral control, social impression, and subjective norms - provided a more comprehensive framework for understanding tourist decision-making. The study's ARIMAX model showed superior performance to traditional ARIMA, ARDL, and VAR models, particularly for short-term predictions, though accuracy declined for longer-term forecasts.

2.6 Resource management in tourist cities

Analysis of resource management research in tourist cities reveals varying approaches to balancing tourism growth with sustainable development. Critical examination of recent studies demonstrates an evolution from traditional management methods to data-driven decision-making.

The Gansu Province study [34] provided valuable insights through its analysis of four resource-dependent cities: Wuwei, Pingliang, Baiyin, and Zhangye. Using coupling coordination models and entropy weight methods, their findings demonstrated how different development stages affect tourism sustainability. Notably, Zhangye's regenerative approach achieved the highest coordination level, while Baiyin's declining status highlighted the risks of unbalanced development. However, the study lacked real-time data integration for dynamic resource allocation.

Research in Samosir, Lake Toba [35] advanced the field by implementing system dynamics and resource-based view (RBV) frameworks. Their identification of three systemic issues - quality gaps, infrastructure limitations, and lake pollution - demonstrated the importance of holistic analysis in tourism management. The study's emphasis on leveraging unique local resources for sustainable competitiveness provided valuable insights. However, it was limited by its reliance on static assessment methods.

The Bali study [36] contributed significant understanding through its analysis of two tourism villages - Taro and Munduk. This research revealed the effectiveness of integrating community involvement, customary practices, and storytelling in resource management, all guided by the Tri Hita Karana philosophy. While successful in demonstrating the value of cultural integration, the study lacked quantitative metrics for resource allocation efficiency.

This review highlights three key research gaps in smart tourism management. First, most studies focus on either data collection or analysis in isolation. Second, frameworks balancing data utility with user privacy remain limited. Third, border city tourism management research lacks comprehensive analysis of tourist behavior.

Our research addresses these gaps through:

1. Integration of spatial-temporal analysis with privacy-preserving techniques
2. Extended temporal analysis covering both regular periods and special events
3. Implementation of data-driven resource allocation methods for border city contexts

Building on previous work [15], this research expands both its scope and sophistication through enhanced data collection, improved privacy protection, and comprehensive analysis frameworks.

3. Materials and methods

This section outlines the materials and methods employed in the current research, organized into five main components: the case study in Sri Chiang Mai Smart City, data collection methods through Wi-Fi infrastructure, details of the public Wi-Fi access logs dataset, analysis methods for tourist behavior patterns, and tourism management system design methods.

3.1 Case study in Sri Chaing Mai Smart City

The Sri Chiang Mai Municipality, a popular district within the Nong Khai Province of Thailand, was selected as a case study site for four strategic reasons. First, its location adjacent to the picturesque Mekong River, directly across from Vientiane, the capital city of Laos, provides a unique opportunity to study border city tourism dynamics. Second, the area features diverse attractions including temples, cultural landmarks, and recreational spaces that attract both domestic and international tourists. Third, the municipality's regular cultural events, such as the Songkran and longboat racing festivals, offer opportunities to study temporal variations in tourist behavior. Fourth, the existing smart city infrastructure and municipal support enables comprehensive data collection and analysis.

The selection of access point locations was driven by four key factors that support effective tourist behavior analysis:

1. Tourist Flow Coverage: Strategic placement along the 4-kilometer riverbank covers major tourist routes
2. Activity Zone Monitoring: Locations near cultural sites and event spaces enable tracking of both regular and special event activities
3. Infrastructure Accessibility: The selected sites have reliable power supplies and network connectivity
4. Privacy Considerations: Placement was optimized to maintain signal strength while respecting private spaces

The Wi-Fi installation project was implemented in two phases to optimize coverage and improve tourist accessibility. Before using the free Wi-Fi service, users must obtain explicit permission through a consent process.

The initial phase began in May 2022 with the deployment of nine access points (APs) along a 4-kilometer stretch of the Mekong riverbank. This area is a major tourist attraction featuring a designated path for running and cycling, while also providing substantial public space for tourist activities and community events. From September 23 to September 29, 2022, the municipality upgraded all nine APs to enhance the Wi-Fi signal strength and increase capacity for the growing number of tourists and local users at each location.

In the second phase, starting in January 2023, the municipality collaborated with NT to optimize AP locations based on tourist movement patterns and usage data collected from May to December 2022. The APs of two locations with low tourist traffic were relocated: the Klong-soi-1 Bridge site was moved to an exercise area at Soi-19, changing Site No. 1 to Site No. 10. Additionally, the AP at the Village Fund Building site was relocated to a community area at Soi-16, changing Site No. 2 to Site No. 11. In Figure 1, the gray dots represent the original sites that were removed and no longer provide service, while the current operational sites serving tourists and locals are marked in green.

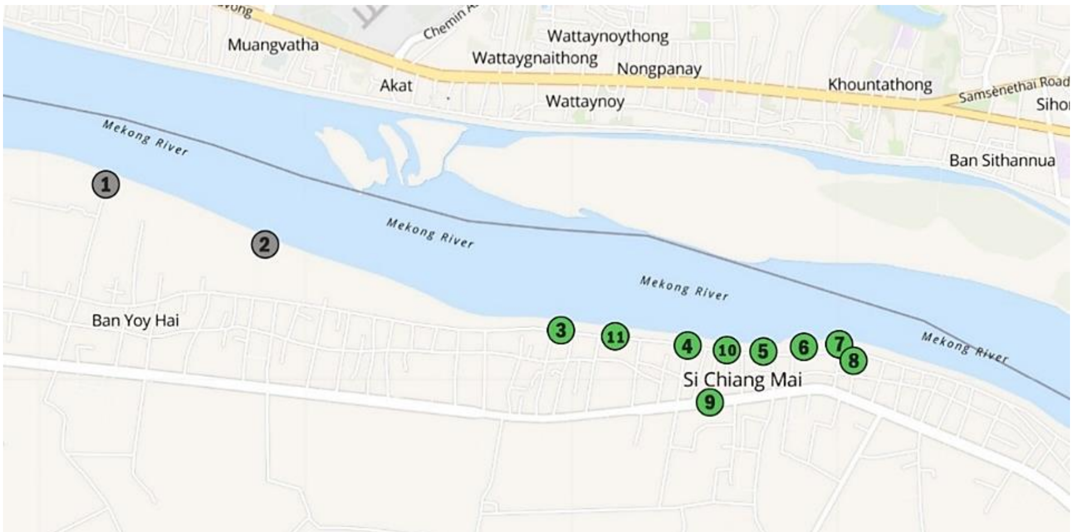

Figure 1 Location of public WI-FI hotspots in Sri Chiang Mai Smart City phase II. [Map source: https://www.openstreetmap.org/]

From May 2022 to December 2023, NT gathered minute-by-minute data to analyze tourist behavior and local usage patterns at each location using the APs as sensors. The system was designed to differentiate between tourist and local resident usage through connection patterns and duration of stay. To protect privacy, all personal data, including MAC addresses, underwent encryption, reducing the risk of individual identification. This study received an exemption from the Khon Kaen University Ethics Committee for Human Research and obtained network access certification from NT.

Comprehensive insights into the environmental factors pertaining to regions offering free Wi-Fi services were obtained, with annual activities organized at Wi-Fi installation sites overseen by the Sri Chiang Mai Municipality, from May 2022 to December 2023, as depicted in Table 1.

Table 1 Descriptions of the environment and its activities for each location

No.	Site Name	Surrounding environment and activity
1	Klong-soi-1 Bridge	The established area has been used for agricultural purposes, specifically for planting seasonal crops such as tomatoes and corn. It is the endpoint of the lane for running and cycling.
2	Village Fund Building	This is a small structure where community members' agricultural products are sold. The lane in front of this site is used for running and cycling.
3	In front of Chang Phueak Temple	This is a Buddhist temple with beautiful architecture, including the temple entrance and the chapel. There is a running and cycling path in front of the temple.
4	In front of Klang Temple	The structure is a Buddhist temple with exquisite architecture at its entrance, chapel, and golden pagoda. A path for running and cycling passes in front of the temple.
5	Sri Chiang Mai Community Health Park	The gardens provide an ideal environment for people to relax and enjoy a variety of food and drinks. The neighborhood elementary school is conveniently located adjacent to the gardens for easy access. There is a path where one can run or bicycle.
6	Naga Courtyard	This area establishes itself as a notable landmark that the entire town values thanks to its impressive Naga statues and captivating gardens. Moreover, a diverse range of food options and beverages are available, adding to its allure. There is a path where one can run or bicycle.
7	In front of Had Pathum Temple	The temple serves as the primary Buddhist symbol and contains magnificent white statues of Buddha and Naga. A spacious courtyard adjacent to the temple serves as a lively venue for events, attracting both locals and tourists.
8	Boeng Wiang Courtyard	This expansive courtyard serves as a versatile location for a variety of events, including annual gatherings and major citywide celebrations. Additionally, it is a charming location for evening activities and social gatherings. It is utilized for recurring cultural celebrations, such as Songkran festival, which takes place from the 13th to the 15th of April, and the longboat racing festival which occurs in October.
9	In front of Fresh Market	This is the only venue in town that provides raw ingredients for cooking, and it does so every morning and afternoon.
10	Soi-19	Evening dance aerobics sessions take place in the area located between Sri Chiang Mai Community Health Park and in front of Klang Temple. There is a path where one can run or bicycle.
11	Soi-16	This is located between the front of Klang Temple and the front of Chang Phueak Temple, next to a community area and a fried clam vendor. There is a path where one can run or bicycle.

3.2 Data collection methods

This research implements a systematic data collection and processing approach, as shown in Figure 2. The process flow has four main stages: raw data collection for gathering Wi-Fi connection data, data processing for cleaning and standardization, privacy protection for data anonymization, and data storage for organizing the processed data. Each stage ensures data quality while maintaining user privacy throughout the collection process.

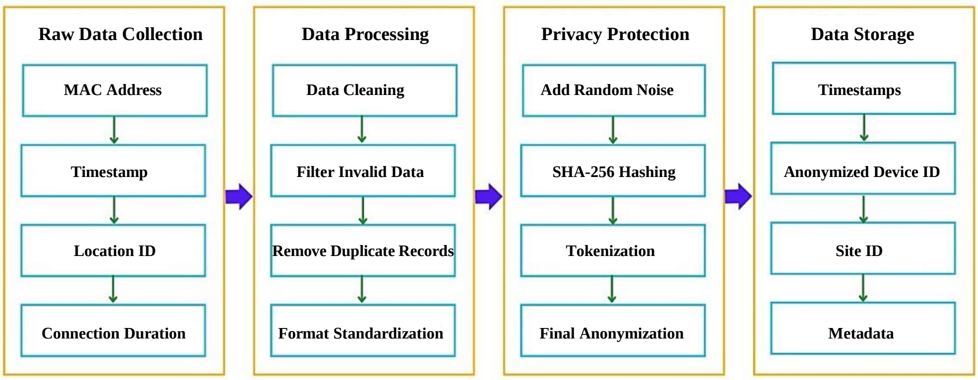

Figure 2 Data collection and privacy protection workflow

3.2.1 Raw data collection

In the initial stage, essential information about devices connecting to the public Wi-Fi network was captured, accumulating 72 million connection records from May 2022 to December 2023. MAC addresses serve as unique device identifiers for tracking tourist movements. Timestamps record connection events enabling temporal analysis of space utilization. Location IDs identify connection points among 11 access points, supporting spatial distribution analysis. Connection duration measurements show how long tourists stay at each location. This raw data forms the foundation for subsequent processing and analysis stages.

3.2.2 Data processing

The data processing phase, handled through the NT secure data center, implements four cleaning procedures to ensure data quality. Data cleaning filters incomplete records from nine access points during the first phase (May-December 2022) and from relocated points in the second phase (January-December 2023). Invalid data removal eliminates records during system maintenance (September 23-29, 2022) and power outages. Duplicate record management prevents double-counting when tourists move between access points, particularly in high-traffic areas like Sri Chiang Mai Community Health Park and Naga Courtyard. Format standardization unifies timestamps and location identifiers across all 11 sites for consistent analysis.

From 72 million connection records, the cleaning process removed 1.66 million records (2.3%). These included incomplete connections after 5 seconds (864,000 records, 1.2%), device handover duplicates (576,000 records, 0.8%), and maintenance period data (216,000 records, 0.3%). Removal of these problematic records enhanced data quality, especially in areas with frequent connection handovers.

3.2.3 Privacy protection

The privacy protection phase, implemented through the NT secure data center, employs four security layers that balance data utility with privacy. Random noise addition introduces controlled variations to MAC addresses, maintaining device recognition within daily periods while preventing direct identification. SHA-256 hashing creates fixed-length values that ensure original identifiers cannot be recovered. Tokenization converts hash values into simplified tokens through consistent mapping, preserving analytical relationships while breaking connections to original identifiers. Final anonymization creates unique identifiers that separate analytical data from personal information.

The system maintains accuracy through multiple control measures. Random noise is calibrated to minimize impact on statistical analysis while maintaining temporal accuracy for monitoring tourist behavior patterns. Spatial precision is kept within access point coverage zones, and device identification accuracy is preserved through consistent hashing algorithms.

3.2.4 Data storage

The data storage phase organizes processed data through the NT secure data center using four structured components. Timestamps store temporal information across three formats: internet_access_logs for access events, monthly CSV files, and a consolidated study period file. The device_info table maintains anonymized device identifiers and hashedmac values with restricted access protocols. Access_points_locations serves as a static lookup table linking site IDs with geographic coordinates for all 11 Wi-Fi points. Metadata supports behavioral analysis during regular periods and special events.

Further details about the database structure and organization appear in Section 3.3, which describes the public Wi-Fi access logs dataset.

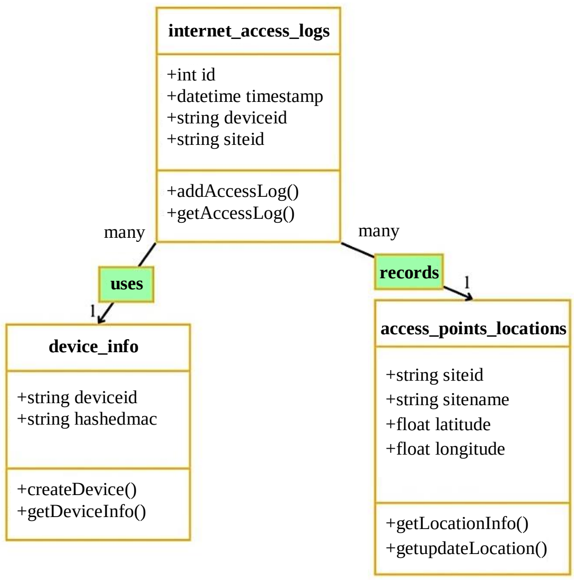

Figure 3 Class diagram for Wi-Fi access logs database

3.3 Detail of Public Wi-Fi access logs dataset

The database management system consists of three main classes with their relationships shown in Figure 3. The class diagram represents the database structure, where internet_access_logs serves as the main class that records tourist Wi-Fi access events. This class uses device identification from the device_info class and records connection locations through the access_points_locations class. Each internet_access_logs entry can be linked to one device and one location, while a single device or location can have many access log entries, as shown by the "many-to-one" relationships in the diagram.

The database structure consists of three main components. First, Tables 2-4 define the core structures: internet_access_logs with user connection details, access_points_locations with site coordinates, and device_info with device identifiers. Second, Tables 5-7 provide sample data for each corresponding structure. Third, Tables 8-10 contain the supporting documentation, including monthly data files, consolidated records, and access point details. This organization ensures efficient data management while maintaining privacy protection.

Table 2 Attributes and descriptions of the internet_access_logs table

Attribute	Description
id	Identifier of the internet_access_logs entry
timestamp	Timestamp of the entry
deviceid	Identifier of the device that accessed the internet
siteid	Identifier of the site or location where the access point is installed

Table 3 Attributes and descriptions of the access_points_locations table

Attribute	Description
siteid	Identifier of the site or location where the access point is installed
sitename	Name of the location
latitude	Latitude coordinate of the location
longitude	Longitude coordinate of the location

Table 4 Attributes and descriptions of the device_info table

Attribute	Description
deviceid	Identifier of the device that accessed the internet
hashedmac	MAC address hashed utilizing the SHA-256 algorithm

Table 5 A sample data snippet of the internet_access_logs table

id	timestamp	deviceid	siteid
1	21/12/2022 0:00	2056	7
2	21/12/2022 0:00	3548	7
…	…	…	…
N	21/12/2022 0:00	10486	11

Table 6 Data entry in the access_points_locations table

siteid	sitename	latitude	longitude
1	Klong-soi-1 Bridge	17.965309	102.560149
2	Village Fund building	17.962984	102.566729
3	In front of Chang Phueak Temple	17.959747	102.578957
4	In front of Klang Temple	17.959126	102.584170
5	Sri Chiang Mai Community Health Park	17.958903	102.587340
6	Naga Courtyard	17.959001	102.588941
7	In front of Had Pathum Temple	17.959003	102.590498
8	Boeng Wiang Courtyard	17.958664	102.590896
9	In front of Fresh Market	17.956875	102.585087
10	Soi-19	17.958885	102.585845
11	Soi-16	17.959502	102.581291

Table 7 A sample data snippet of the device_info table

deviceid	hashedmac
1	9b3708025e7e1dcf07fa8139e7c5f0385ff585c507eb582a9d6259c31077ec80
2	abdfc73c7bc0bd34860eb7522713fd5539f2208985f412badbdb42cdfeb192cd
…	…
N	493cf10794abd6336857a49882ef4a650396d9654750bea98fbc5c74d0ef656f

Table 8 Monthly data files from May 2022 to December 2023

Name	Description
2022_m05_client_wifi.csv	The CSV file for the month of May 2022
2022_m06_client_wifi.csv	The CSV file for the month of June 2022
…	…
2023_m12_client_wifi.csv	The CSV file for the month of December 2023.

Table 9 A single file that consolidates data from May 2022 to December 2023

Name	Description
total_2022_m05_to_2023_m12_client_wifi.csv	A CSV file that consolidates monthly data from May 2022 to December 2023.

Table 10 Access point locations documentation file

Name	Description
access_points_locations.csv	A CSV file documenting the locations of the access points.

3.4 Analysis methods

This research implements spatial and temporal analysis methods offering key advantages over alternatives. Our Wi-Fi based tracking combines wide coverage, indoor-outdoor functionality, and non-intrusive monitoring, despite potential signal interference and privacy concerns. In comparison, Bluetooth beacons have limited range, GPS performs poorly in urban areas, camera systems are weather-dependent, and mobile apps require active user participation.

The temporal analysis using minute-by-minute data collection provides higher granularity than traditional methods, better accuracy than manual counting, and enables real-time pattern detection. These methods specifically address border city tourism challenges, enabling a nuanced understanding of visitor patterns during both regular periods and cultural events.

3.4.1 Quantitative analysis methods

Statistical analysis examines tourist numbers through Wi-Fi connection data. Monthly tourist calculations use unique anonymized device IDs from all access points. This analysis includes mean value computation for monthly tourist numbers and standard deviation analysis to measure monthly variations.

3.4.2 Temporal analysis methods

Time pattern analysis focuses on major cultural events in Sri Chiang Mai Smart City. Hourly analysis identifies peak tourist activities each day, while monthly analysis examines tourist patterns during Songkran in April and the longboat racing festival in October.

3.4.3 Spatial analysis methods

Tourist distribution analysis examines connection densities across all 11 Wi-Fi access points in Sri Chiang Mai Smart City. Bar chart visualization techniques are applied to compare total connections between all locations along the Mekong River area to understand overall tourist distribution patterns.

3.4.4 Pattern analysis methods

Tourist behavior analysis during special events examines three aspects: festival impact patterns during the Songkran festival, usage comparisons between regular and special event days including the longboat racing festival, and location preferences during various activities.

3.4.5 Behavioral prediction methods

Tourist behavior prediction utilizes pattern analysis. The method analyzes three predictive aspects: festival-based predictions drawing from annual cultural events, time-based predictions analyzing daily connection trends, and location-based predictions examining tourist distribution across access points. These patterns support future event planning and resource allocation in Sri Chiang Mai Smart City.

3.5 Tourism management system design methods

The tourist behavior analysis system implements a comprehensive design methodology to support smart tourism management in Sri Chiang Mai Smart City. The system design consists of three main components: system architecture design for defining structural relationships, system development process for implementing data processing capabilities, and system testing for ensuring reliability and data protection.

3.5.1 System architecture design

The tourist behavior analysis system architecture implements a three-layer design as illustrated in Figure 4. This layered architecture ensures the separation of concerns while maintaining efficient data flow and processing:

1. The data layer comprises three main classes. An internet_access_logs class manages tourist connection records with key attributes including deviceid and siteid for identifying visitor activities. The access_points_locations class handles spatial data from 11 Wi-Fi access points along the Mekong River area. A device_info class maintains anonymized MAC addresses for privacy protection.

2. The processing layer handles data collection and analysis functions. Wi-Fi connection data passes through privacy protection using SHA-256 hashing and tokenization. The system implements data-cleaning procedures and standardizes formats for analysis. This layer managed the processing of 72 million connection records through Apache Spark.

3. The application layer delivers tourist behavior analysis capabilities. It enables monitoring of tourist densities across locations, analyzes temporal patterns during festivals, and examines spatial distributions of visitors. The relationships between classes support comprehensive analysis while maintaining data privacy.

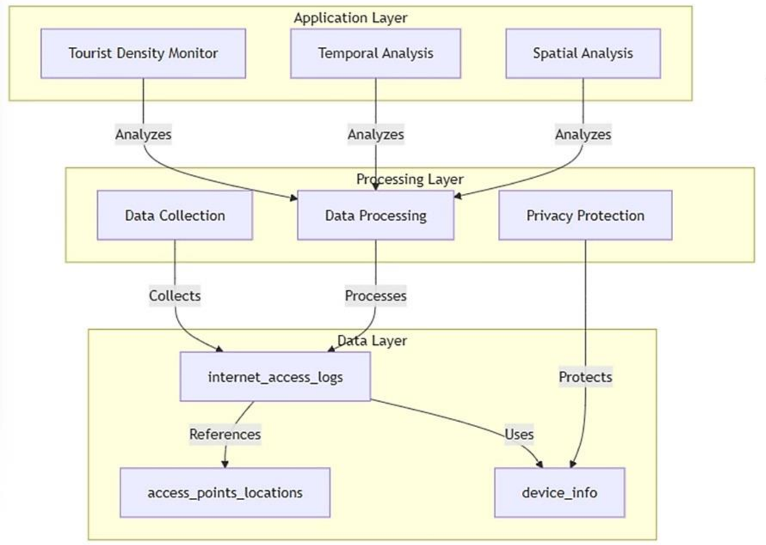

Figure 4 Three-Layer architecture of the tourist behavior analysis system

3.5.2 System development process

This research implements development in four main steps using Apache Spark for large-scale data processing and Python for analysis:

1. Data collection implementation develops data capture mechanisms across 11 Wi-Fi access points through virtual Local Area Network segmentation. The process includes tourist device tracking, timestamp recording, and site identification systems. Each access point installation follows the Sri Chiang Mai Municipality's specifications for smart city development.

2. The data processing pipeline implements privacy protection and data standardization. This system processes 72 million connection records through SHA-256 hashing for MAC address anonymization. Data cleaning procedures remove incomplete records and standardize formats for analysis.

3. Analysis module development creates functions for tourist behavior analysis. The module calculates tourist density, identifies temporal patterns during festivals, and analyzes spatial distributions. These components support both daily monitoring and special event analysis.

4. Database management implements three core tables: internet_access_logs for connection records, access_point_locations for spatial data, and device_info for anonymized identifiers. The database structure enables efficient data retrieval while maintaining privacy protection.

3.5.3 System testing

This research implements testing procedures in three key areas to ensure system reliability and data protection:

1. Data validation testing ensures data accuracy and completeness throughout the system. The validation process verified MAC address anonymization accuracy across all access points from May 2022 to December 2023. Location data testing confirms precise tracking of tourist movements across 11 sites, with specific validation for relocated access points at Soi-19 and Soi-16. Connection record validation ensures accurate timestamp recording for temporal analysis.

2. Performance testing evaluates system capabilities in handling large-scale data processing. The testing validates Apache Spark processing efficiency with 72 million records and assesses real-time data collection performance across all access points. Database query testing ensures efficient data retrieval for tourist behavior analysis during both regular periods and festival events.

3. Privacy protection testing verifies the security of tourist data throughout the system. Testing procedures validate SHA-256 hashing implementation and data anonymization effectiveness. The process ensures compliance with Khon Kaen University Ethics Committee requirements while maintaining data utility for analysis.

4. Results

The public Wi-Fi infrastructure in Sri Chiang Mai Smart City provides comprehensive data from May 2022 to December 2023. The study results are analyzed in six parts: quantitative analysis examining tourist statistics, temporal analysis investigating time-based patterns, spatial analysis exploring location-based trends, pattern analysis revealing behavioral characteristics, behavioral prediction showing future trends, and implementation results demonstrating practical applications. These analyses enhance understanding of dynamic patterns and provide crucial insights for urban planning and community engagement.

4.1 Quantitative analysis results

As shown in Figures 5 and 6, statistical analysis of Wi-Fi connection data demonstrates both monthly and daily tourist patterns. Detailed analysis of these patterns reveals the following characteristics.

In monthly analysis from May 2022 to December 2023, the average monthly count of tourists is 3,403.40, with a 95% confidence interval ranging from 2,892.52 to 3,914.28, confirming the robustness of the results. The median value is 3,267.00, indicating a slight positive skew, likely influenced by high tourist volumes during festival months. The variance measures 1,358,825.41, and the standard deviation is 1,165.69 tourists, reflecting substantial monthly variations in tourist flows.

Tourist numbers grew significantly over the study period, progressing from an average of 2,282 tourists per month in 2022 to 4,241 tourists per month in 2023, representing an 85.8% year-over-year increase. Monthly counts registers each visitor once regardless of how many days they visited during that month. This growth aligns with the research objective of analyzing trends to inform resource allocation. Notably, June 2022 recorded the lowest tourist volume, 1,223, while October 2023 saw a peak of 5,114 tourists, likely driven by cultural events such as the longboat racing festival. These findings highlight the influence of seasonal events on tourism patterns, underscoring the need for event-specific resource planning.

Daily analysis reveals consistent weekly trends with an average daily count of 10,112.43 tourists and a 95% confidence interval of 9,603.93 to 10,620.93, demonstrating the reliability of the observed data. The daily counts accumulate repeated visits from the same individuals across different days. The median value is 9,894.00, while the variance and standard deviation are 471,176.27 and 686.42 tourists, respectively, indicating relatively stable daily fluctuations.

A clear weekly pattern emerges, with weekday tourist numbers ranging from 9,500 to 9,900, and significant increases during weekends exceeding 10,300 tourists, peaking at 11,424 tourists on Sundays. This predictable weekend surge highlights opportunities for optimizing service schedules, such as increasing staff availability and enhancing infrastructure during peak periods.

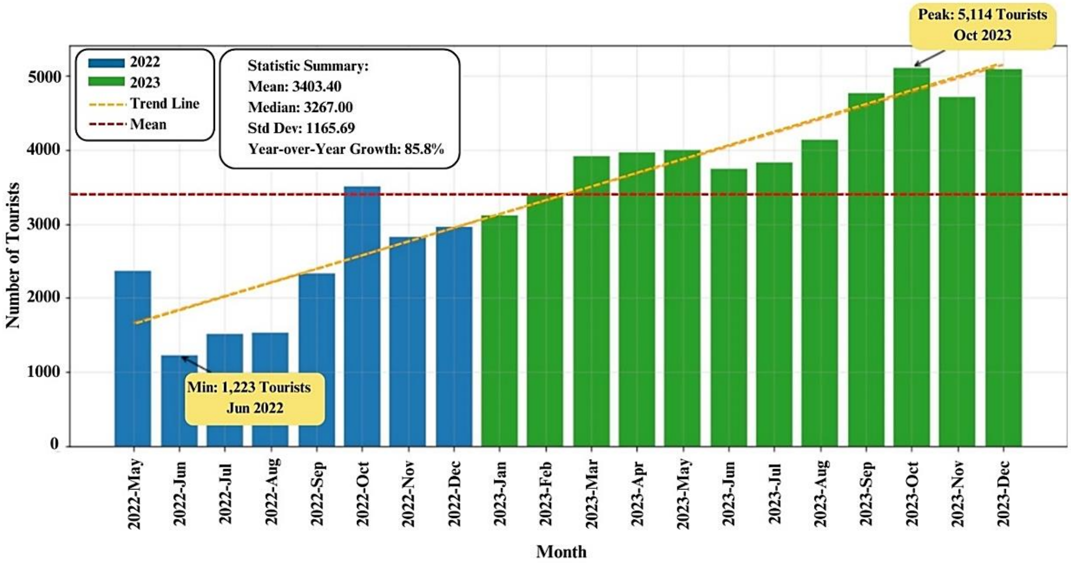

Figure 5 Monthly tourist analysis through Wi-Fi connection data in Sri Chiang Mai Smart City from May 2022 to December 2023

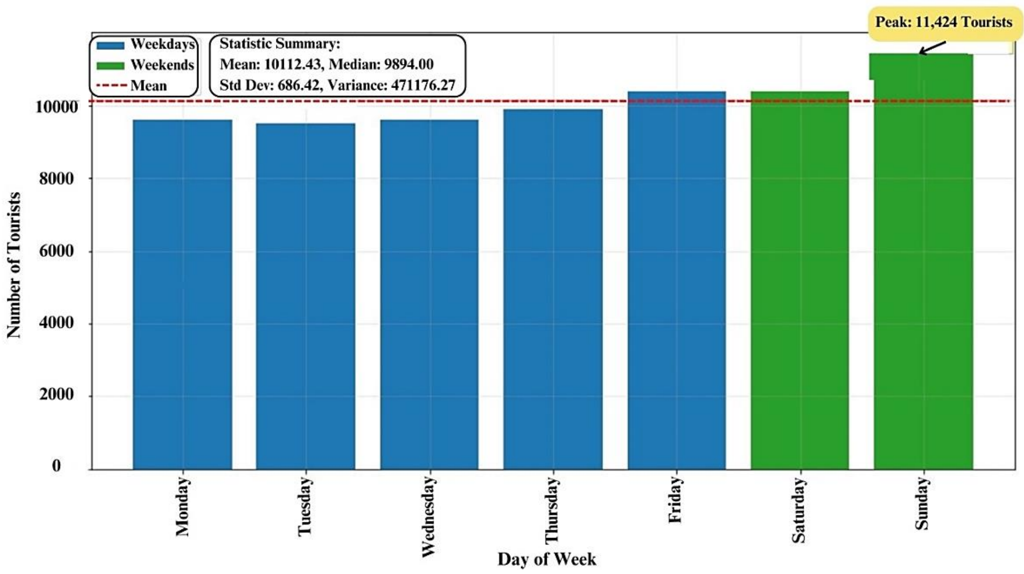

Figure 6 Weekly tourist analysis by days through Wi-Fi connection data in Sri Chiang Mai Smart City

These results validate the effectiveness of Wi-Fi data in identifying temporal patterns. The observed growth in monthly tourist numbers aligns with the objective of analyzing trends to support infrastructure planning. The weekend surge effect emphasizes the importance of real-time data in managing dynamic tourist flows, particularly during high-demand periods. These findings also align with prior research [20, 21] that identifies cultural events as key drivers of tourist behavior, though the observed strong evening preference (17:00–21:00) offers a unique contribution to the literature.

4.2 Temporal analysis results

Analysis of Wi-Fi connection data reveals distinct temporal patterns between event and non-event locations, as shown in Figures 7 and 8. Tourist numbers showed dramatic increases during cultural events, with a 127.87% increase during the Songkran festival from April 13 to 15 compared to regular days in the same month at the Boeng Wiang Courtyard. At this main event venue, tourist numbers significantly increased during the festival period, with April 14 recording the highest attendance. The 95% confidence interval for tourist numbers during the festival period ranges from 312.45 to 336.89 visitors, confirming the reliability of the observed surge. Hourly analysis of April 14 reveals peak tourist activities between 17:00 to 21:00, coinciding with event activities such as beauty contests and concerts. These findings emphasize the importance of allocating additional resources, such as staff and security, during peak hours to enhance visitor experiences and ensure crowd safety.

In contrast, the front of Chang Phueak Temple maintained stable visitor levels throughout both the festival period and during April 14, with a smaller fluctuation range (95% confidence interval: 65.33 to 78.67 visitors). This stability suggests that non-event locations experience consistent attraction patterns regardless of cultural events, underscoring the complementary role of such locations in distributing tourist activity.

These findings align with prior studies [20, 23] that highlight the impact of cultural events on tourist behavior. However, the strong evening preference (17:00–21:00) contrasts with daytime patterns observed in urban contexts [21], emphasizing the need for localized strategies. These results support the research objective of optimizing resource allocation during high- demand periods.

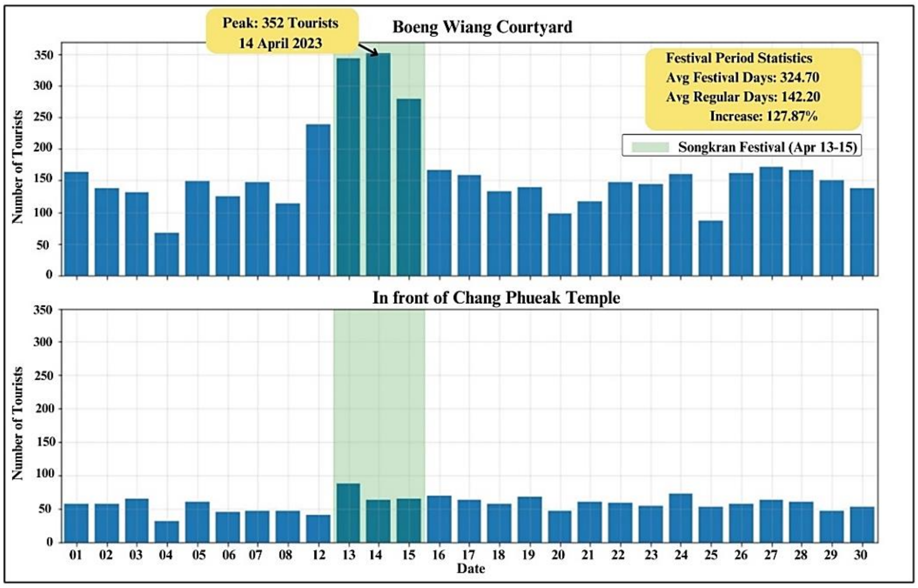

Figure 7 Daily tourist comparison between event (Boeng Wiang Courtyard) and non-event locations (Chang Phueak Temple) during the Songkran Festival, April 2023

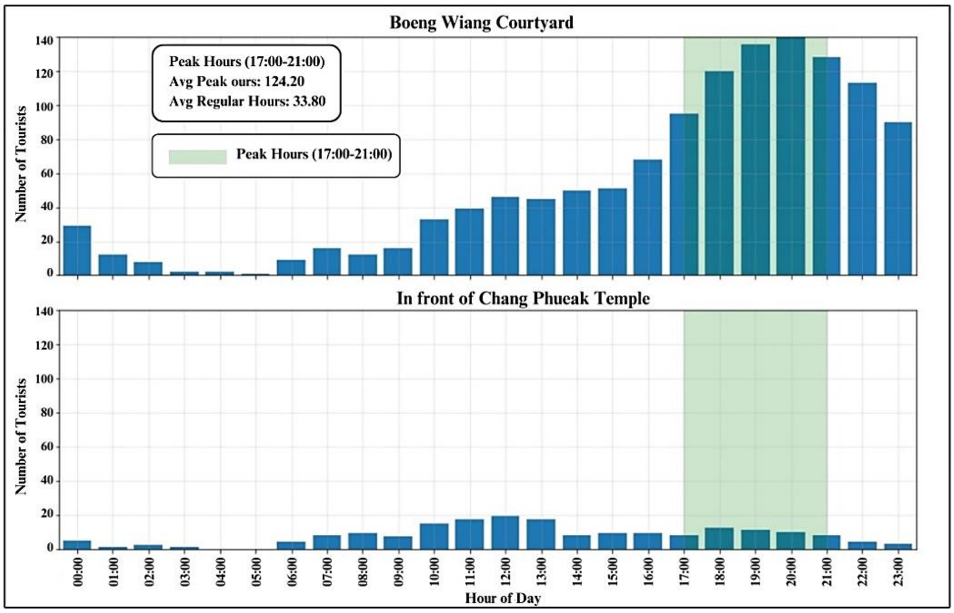

Figure 8 Hourly tourist analysis between event and non-event locations on April 14, 2023 during peak Songkran Festival activities

4.3 Spatial analysis results

Line graph analysis showing tourist distribution across 11 Wi-Fi access points from May 2022 to December 2023 reveals spatial patterns across different timeframes, as shown in Figures 9 and 10. Four locations consistently showed higher visitor concentrations: Sri Chiang Mai Community Health Park, Naga Courtyard, In front of Had Pathum Temple, and the Boeng Wiang Courtyard. Monthly analysis in Figure 9 identifies these main high-traffic locations maintaining stable visitation patterns throughout the study period. Daily analysis shown in Figure 10 demonstrates these same locations maintaining consistently high tourist numbers throughout the week, with the fresh market showing regular activity patterns independent of weekday variations. This spatial distribution aligns with the locations' roles as primary cultural and recreational areas, supporting our research objective of optimizing tourism resource allocation and infrastructure development based on actual usage patterns.

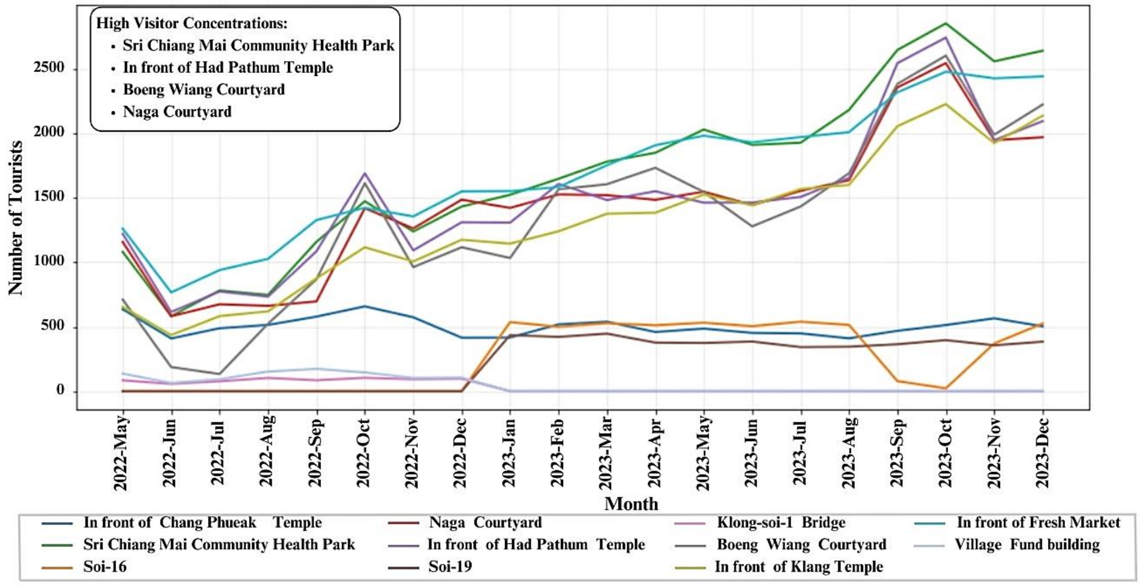

Figure 9 Tourist distribution analysis across 11 Wi-Fi access points in Sri Chiang Mai Smart City from May 2022 to December 2023

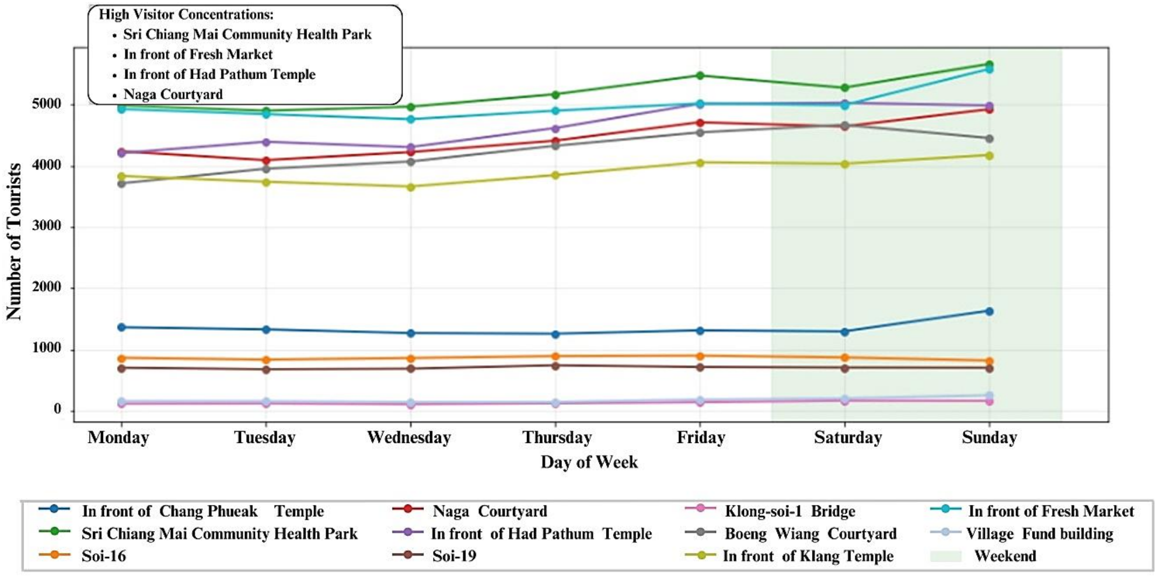

Figure 10 Daily tourist distribution patterns across key locations in Sri Chiang Mai Smart City

4.4 Pattern analysis and behavior prediction

Analysis of Wi-Fi usage data from May 2022 to December 2023 reveals consistent behavioral patterns and enables predictive insights. Festival events demonstrate significant impact on tourist behavior, with the Songkran festival exhibiting a 127.87% increase in visitors at the Boeng Wiang Courtyard compared to regular periods. Peak usage consistently occurs between 17:00-21:00. This is especially pronounced during festivals and weekends.

Location preferences show clear patterns across the city. Four locations maintain consistently higher visitor concentrations: Sri Chiang Mai Community Health Park, Naga Courtyard, the front of the Had Pathum Temple, and the Boeng Wiang Courtyard. These patterns vary between event and non-event periods, with the Boeng Wiang Courtyard showing dramatic increases during festivals while Sri Chiang Mai Community Health Park maintains stable visitation throughout the year.

These established patterns enable reliable predictions of tourist behavior. Annual cultural events, particularly the Songkran festival in April and longboat racing festival in October, show predictable visitor surges. Daily patterns indicate optimal periods for resource deployment, with evening hours requiring enhanced service levels. Location-based predictions help optimize resource allocation, particularly during high-traffic periods at primary attractions.

4.5 Implementation results

The analysis demonstrates practical applications across three key areas. For tourism management, the system enables evidence-based resource allocation and service scheduling according to identified patterns. Stakeholder benefits include improved decision-making for government agencies and optimized operations for tourism businesses. Smart city development advances through infrastructure optimization, demonstrated by the successful relocation of access points to higher-traffic areas at Soi-19 and Soi-16.

Usage patterns inform facility development, particularly in high-traffic areas. The event management system benefits from data-driven capacity planning, while service enhancement strategies focus on peak evening hours from 17:00-21:00. These implementations directly support the research objective of developing effective smart tourism management solutions.

5. Discussion

This research demonstrates understanding of tourist behavior through public Wi-Fi data analysis across three key dimensions. Its findings both align with and differ from previous smart city implementations. First, behavioral analysis reveals distinct usage patterns with peak activities during evening hours (17:00-21:00) and during major festivals, particularly at the Boeng Wiang Courtyard during the Songkran festival. While this evening preference contrasts with studies in Turin [17] and the Madeira Islands [18], where peak tourist activities occurred during daytime, it aligns with findings from Singapore's Marina Bay study [20] where evening events significantly influenced tourist distribution. The data enables urban planners to allocate resources and services according to demand patterns specific to border city contexts. Second, the tourist behavior analysis system development incorporates privacy protection through MAC address encryption using SHA-256 and tokenization, similar to the approach used in Lisbon's hotspot study [29]. However, while the Lisbon system achieved only 14% tracking accuracy for outdoor locations, our system maintained consistent tracking capabilities across both indoor and outdoor spaces. The system successfully processed over 72 million connection records throughout the 20-month study period, demonstrating robust scalability comparable to Florence's comprehensive study [30] which analyzed 56 million connections. Third, our findings on spatial distribution patterns align with research from the Shichahai Scenic Area [27] where tourist concentrations correlated strongly with cultural and historical sites. The successful relocation of access points from low-traffic areas to more active zones demonstrates the system's effectiveness in supporting dynamic urban management, similar to the adaptive approach used in Bologna's university area study [19].

Several limitations should be acknowledged. Regarding data coverage, the study is limited to tourists using public Wi-Fi, excluding visitors using mobile data or no digital devices, with potential underrepresentation of international tourists using roaming data. In terms of spatial constraints, the study area is limited to locations along the Mekong River and may not capture complete tourist movement patterns, with potential gaps in coverage between Wi-Fi access points. Technical limitations include signal interference in crowded areas, weather effects on outdoor Wi-Fi performance, and dependency on a continuous power supply. Additionally, the data processing and privacy protection measures introduced minor trade-offs between accuracy and privacy. While the cleaning process removed 2.3% of potentially useful partial connections, it significantly improved dataset reliability. Similarly, addition of random noise slightly reduced temporal precision but enhanced privacy protection. These impacts were deemed acceptable given the large sample size of 72 million records and maintained statistical significance of the results. The confidence intervals reported in the results section reflect these considerations, providing transparent bounds for the accuracy of our findings. These findings have significant implications for future infrastructure investments, especially in developing public spaces along the Mekong River with tourism potential, and planning community engagement strategies that align with observed usage patterns.

6. Conclusion

This research achieved its objective of developing a tourist behavior analysis system using public Wi-Fi data to enhance smart tourism management. The system offers several advantages over traditional monitoring methods: continuous data collection without tourist disruption, deep insights into temporal and spatial patterns, and support for data-driven decision making. The practical implications are particularly significant for policymakers and urban planners through resource allocation that enables evidence-based distribution of tourism resources, event planning that supports optimal scheduling and location selection for cultural events, infrastructure development that guides strategic placement of tourist facilities, and community impact that helps balance tourist services with local needs.

However, the research has important limitations in its current form. Future research directions should consider data integration by combining Wi-Fi data with mobile phone data for broader coverage, incorporating social media check-ins and reviews, and adding environmental sensors for context-aware analysis. Spatial expansion could include extending coverage to the entire city area, including secondary tourist locations, and analyzing cross-border movement patterns. Analytical advancement might focus on developing AI-based prediction models, implementing real-time analysis capabilities, and creating interactive visualization tools. Environmental factors such as studying weather impact on tourist behavior, analyzing seasonal variations, and assessing environmental sustainability should also be considered. These findings contribute to the growing body of knowledge on smart tourism management while providing practical tools for urban development in border city contexts.

7. Data availability statement

The public Wi-Fi access logs dataset generated using Sri Chiang Mai public Wi-Fi infrastructure provides an insight into minute-by-minute analysis of public space utilization. This data reflects the efforts behind its collection, processing, and storage over the duration of the study. The dataset, available at https://doi.org/10.6084/m9.figshare.23691336.v11, possesses a dynamic and continuously updated nature. Users can expect freshly added data every month, providing them with ongoing insights into recent public space utilization in Sri Chiang Mai.

8. Code availability statement

The Python programming language was used to execute the code within a Jupyter notebook environment. The Python scripts were specifically designed to perform tasks such as data preprocessing, visualization, and public space utilization analysis. The scripts, tailored to specific circumstances, can be accessed in the 'wifi_clientdata' GitHub repository. The resources are available through the GitHub link (https://github.com/thalerngsak-wi/srichiangmaiwifi/blob/main/wifi_client.ipynb).

9. Acknowledgements

The authors thank the Sri Chiang Mai Municipality for offering essential perspectives and guidance, which significantly aided in comprehending the local context and facilitated the fieldwork. Further appreciation is extended to the National Telecom Public Company Limited for granting access to public Wi-Fi infrastructure, ensuring network security, offering data server hosting, and sharing their telecommunications expertise. Lastly, sincere acknowledgment is given to the Center of Multidisciplinary Innovation Network Talent, Faculty of Interdisciplinary Studies, Department of Technology and Engineering, Nong Khai Campus, Khon Kaen University for their consistent provision of resources, facilities, and support throughout the research process.

10. Ethical approval

This study was reviewed and granted an exemption by the Khon Kaen University Ethics Committee for Human Research (KKUEC) in accordance with the committee Exemption Determination Regulations under reference HE653311.

11. References

[1] Tian Z, Liu Y, Wang Y, Wu L. A tourist behavior analysis framework guided by geo-information tupu theory and its application in Dengfeng City, China. ISPRS Int J Geo-Inf. 2022;11(4):250.
[2] Gao Y, Liu S, Wei B, Zhu Z, Wang S. Using Wi-Fi probes to evaluate the spatio-temporal dynamics of tourist preferences in historic districts' public spaces. ISPRS Int J Geo-Inf. 2024;13(7):244.
[3] Wang L. Enhancing tourism management through big data: design and implementation of an integrated information system. Heliyon. 2024;10(20):e38256.
[4] Vega-Barbas M, Álvarez-Campana M, Rivera D, Sanz M, Berrocal J. AFOROS: A low-cost Wi-Fi-Based monitoring system for estimating occupancy of public spaces. Sensors. 2021;21(11):3863.
[5] Alorda-Ladaria B, Ruiz-Pérez M, Ramos V. A near real-time monitoring system using public WI-FI data to evaluate COVID-19 social distance measures. Electronics. 2022;11(18):2897.
[6] Salas P, Ramos V, Ruiz-Pérez M, Alorda-Ladaria B. Methodological proposal for the analysis of urban mobility using Wi-Fi data and artificial intelligence techniques: the case of Palma. Electronics. 2023;12(3):504.
[7] Zanlungo F, Feliciani C, Yücel Z, Jia X, Nishinari K, Kanda T. A pure number to assess "congestion" in pedestrian crowds. Transp Res C: Emerg Technol. 2023;148:104041.
[8] Du R, Santi P, Xiao M, Vasilakos AV, Fischione C. The sensable city: a survey on the deployment and management for smart city monitoring. IEEE Commun Surv Tutor. 2019;21(2):1533-60.
[9] Zhang L, Cui W, Li B, Chen Z, Wu M, Gee TS. Privacy-preserving cross-environment human activity recognition. IEEE Trans Cybern. 2023;53(3):1765-75.
[10] Zhao P, Liu W, Zhang G, Li Z, Wang L. Preserving privacy in WiFi localization with plausible dummy locations. IEEE Trans Veh Technol. 2020;69(10):11909-25.
[11] Simončič A, Mohorčič M, Mohorčič M, Hrovat A. Non-intrusive privacy-preserving approach for presence monitoring based on WiFi probe requests. Sensors. 2023;23(5):2588.
[12] Bolleter J. Counter intelligence: evaluating Wi-Fi tracking data for augmenting conventional public space–public life surveys. Aust Plan. 2017;54(2):134-44.
[13] Di Cara N, Zelenka N, Davis O, Haworth C. Using data hazards to support safe and ethical digital footprint research. Int J Popul Data Sci. 2023;8(3):2279.
[14] Kaššaj M, Peráček T. Sustainable connectivity—integration of mobile roaming, WiFi4EU and smart city concept in the European union. Sustainability. 2024;16(2):788.
[15] Wiangwiset T, Surawanitkun C, Wongsinlatam W, Remsungnen T, Siritaratiwat A, Srichan C, et al. Design and implementation of a real-time crowd monitoring system based on public wi-fi infrastructure: a case study on the sri chiang mai smart city. Smart Cities. 2023;6(2):987-1008.
[16] Hu X, Shen P, Shi Y, Zhang Z. Using Wi-Fi probe and location data to analyze the human distribution characteristics of green spaces: a case study of the Yanfu Greenland Park, China. Urban For Urban Green. 2020;54:126733.
[17] Pronello C, Anbarasan D, Spoturno F, Terzolo G. A low-cost automatic people-counting system at bus stops using Wi-Fi probe requests and deep learning. Public Transp. 2024:1-30.
[18] Ribeiro M, Nunes N, Nisi V, Schöning J. Passive Wi-Fi monitoring in the wild: a long-term study across multiple location typologies. Pers Ubiquit Comput. 2022;26(3):505-19.
[19] Boulanger SO, Longo D, Roversi R. Data evidence-based transformative actions in historic urban context—The Bologna university area case study. Smart Cities. 2020;3(4):1448-76.
[20] Zhou Y, Lau BPL, Koh Z, Yuen C, Ng BKK. Understanding crowd behaviors in a social event by passive wifi sensing and data mining. IEEE Internet Things J. 2020;7(5):4442-54.
[21] Li L, Chen X, Li Q, Tan X, Chen J, Wang D, et al. Contextualizing human dynamics: understanding the semantics of movement trajectories with Wi-Fi data. Travel Behav Soc. 2021;25:183-92.
[22] Stanitsa A, Hallett SH, Jude S. Investigating pedestrian behaviour in urban environments: a Wi-Fi tracking and machine learning approach. Multimodal Transportation. 2023;2(1):100049.
[23] Gao Y, Schmöcker JD. Estimation of walking patterns in a touristic area with Wi-Fi packet sensors. Transp Res Part C Emerg Technol. 2021;128:103219.

[24] Fantozzi P, Maccario G, Naldi M. Uncovering tourist visit intentions on social media through sentence transformers. Information. 2024;15(10):603.
[25] Shrestha D, Wenan T, Shrestha D, Rajkarnikar N, Jeong SR. Personalized Tourist recommender system: a data-driven and machine-learning approach. Computation. 2024;12(3):59.
[26] Crivellari A, Beinat E. LSTM-Based deep learning model for predicting individual mobility traces of short-term foreign tourists. Sustainability. 2020;12(1):349.
[27] Li L, Pei Z, Li Q, Hao F, Chen X, Chen J. Identifying tourism attractiveness based on intra-destination tourist behaviour: evidence from Wi-Fi data. Curr Issues Tourism. 2024;27(19):3131-49.
[28] Ruiz-Pérez M, Ramos V, Alorda-Ladaria B. Integrating high-frequency data in a GIS environment for pedestrian congestion monitoring. Inf Process Manage. 2023;60(2):103236.
[29] Claro R, Eisa S, Pardal ML. Lisbon Hotspots: Wi-Fi access point dataset for time-bound location proofs. arXiv:2208.04741:2022.
[30] Bellini P, Cenni D, Nesi P, Paoli I. Wi-Fi based city users' behaviour analysis for smart city. J Vis Lang Comput. 2017;42:31-45.
[31] Li X, Law R, Xie G, Wang S. Review of tourism forecasting research with internet data. Tour Manag. 2021;83:104245.
[32] Zhang X, Cheng M, Wu DC. Daily tourism demand forecasting and tourists' search behavior analysis: a deep learning approach. Int J Mach Learn Cyber. 2024:1-14.
[33] Padhi SS, Pati RK. Quantifying potential tourist behavior in choice of destination using Google Trends. Tour Manag Perspect. 2017;24:34-47.
[34] Deng H, Hu J. A comparative analysis of the coupling and coordination between the ecological civilization construction and tourism development in different types of resource-based cities. J Nat Conserv. 2024;79:126563.
[35] Simarmata E, Kusumastuti R, Wijaya C. Achieving sustainable competitiveness of tourism dynamics with resource-based view. Eur J Innov Manag. 2024;27(3):1001-23.
[36] Rosalina PD, Dupre K, Wang Y, Putra IND, Jin X. Rural tourism resource management strategies: a case study of two tourism villages in Bali. Tour Manag Perspect. 2023;49:101194.