# CHARACTERIZING PHYSICAL SPACE UTILIZATION USING WI-FI NETWORK CONNECTIVITY DATA

Susika CHAIKUL[1] and Santi PHITHAKKITNUKOON[1]
1 Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand;
sunsika_chaikul@cmu.ac.th (S. C.); santi@eng.cmu.ac.th (S. P.)

**ARTICLE HISTORY**
**Received:** 6 January 2025 **Revised:** 20 January 2025 **Published:** 28 January 2025

## ABSTRACT
This paper presents methodologies for analyzing space utilization dynamics and their implications for sustainable building management. Leveraging Wi-Fi connectivity data from a multi-campus university network, we developed a novel framework combining principal component analysis (PCA), eigendecomposition, and X-means clustering to identify and evaluate key features of space utilization: crowdedness, mobility, and connectivity entropy. These metrics provide a comprehensive view of spatial behavior, enabling nuanced insights into how spaces are used over time. The study further examines the relationship between space utilization and electricity consumption, focusing on ten buildings at Chiang Mai University. By correlating Wi-Fi probe data with energy usage patterns, the research highlights the complex interplay between crowdedness, mobility, and entropy with electricity demand. This integrated approach reveals actionable insights into optimizing energy consumption based on real-time spatial behavior. The findings contribute to discussions on sustainable building practices and urban informatics, offering a pathway for improving energy efficiency while enhancing user experience in shared spaces. The methodologies and insights presented in this study serve as a foundation for future research aimed at promoting greener, more resource-efficient urban environments through data-driven decision-making.
**Keywords:** Wi-Fi network, Space utilization, Energy consumption, Space segmentation

**CITATION INFORMATION:** Chaikul, S., & Phithakkitnukoon, S. (2025). Characterizing Physical Space Utilization Using Wi-Fi Network Connectivity Data. *Procedia of Multidisciplinary Research, 3*(1), 11.

## INTRODUCTION

The rapid growth of urban areas has driven cities to adopt smart city frameworks, integrating information and communication technologies (ICT) to address challenges like pollution and transportation inefficiencies (Cho & Choi, 2014; United Nations, 2014). Central to this transformation is the use of sensing technologies, including CCTV, GPS, and wireless networks, to gather data on citizen behavior for informed decision-making and policy development. Opportunistic sensing, such as analyzing Wi-Fi connectivity data initially collected for network monitoring, has gained traction due to its scalability and minimal privacy concerns (Araico, 2017; Shen & Stopher, 2014).

Wi-Fi, a ubiquitous wireless protocol under IEEE 802.11 standards (Lemstra et al., 2010), offers a valuable data source for studying human behavior and space utilization. Previous studies have utilized Wi-Fi data to examine area crowdedness (Araico, 2017), mobility patterns (Sapiezynski et al., 2015; Uras et al., 2020), and occupancy dynamics (Ouf et al., 2017; Sevtsuk et al., 2008). These findings demonstrate Wi-Fi's potential for clustering spaces by usage and modeling movement patterns based on access point connectivity (Calabrese et al., 2010; Kim & Kotz, 2005). By leveraging this data, researchers have explored innovative approaches to spatial planning and building design, offering cost-effective alternatives to traditional methods (Ouf et al., 2017; Sevtsuk et al., 2008).

Buildings, as major contributors to urban energy consumption, represent a critical focus area for sustainable practices (Cao et al., 2016). The interplay between space utilization and energy demand has garnered attention, with studies revealing how occupancy-driven models can optimize heating, cooling, and lighting systems (Erickson et al., 2014). Emerging technologies now allow real-time monitoring of space occupancy and energy consumption, enabling dynamic adjustments to align with usage patterns (Gui et al., 2020; Tien et al., 2022). Despite these advancements, the causal relationship between space utilization and electricity consumption remains underexplored.

This research bridges these gaps by leveraging Wi-Fi connectivity data to investigate space utilization and its correlation with energy consumption. The first study introduces the Xplaces method, a framework employing principal component analysis (PCA) and clustering techniques to analyze features like crowdedness, mobility, and entropy. Using data from a university network, Xplaces offers insights into spatial segmentation and optimal feature selection. The second study extends this work to examine the relationship between space utilization and electricity consumption, identifying real-time indicators of energy demand based on spatial dynamics.

By integrating findings from these studies, this research aims to advance data-driven strategies for optimizing physical spaces and energy use. These contributions offer actionable insights for urban planners, architects, and facility managers, promoting sustainable and efficient practices in urban environments.

## RESEARCH METHODOLOGY

This study utilized Wi-Fi connectivity data from Chiang Mai University's network, encompassing 2,980 access points (APs) across three campuses: Suan Sak (2.93 $km^2$, administrative and academic hub), Suan Dok (0.45 $km^2$, health sciences complex), and Mae Hia (3.50 $km^2$, veterinary and agro-industry faculties), covering a total area of 6.88 $km^2$.

The dataset included connectivity details such as device MAC addresses, RSSI, AP identifiers, timestamps, and geolocation data, sampled every 5 minutes from January 9 to February 3, 2020. This yielded over 133 million records from 291,124 unique devices. APs were categorized into six groups: residences, academic buildings, administrative buildings, service centers, research institutes, and others (e.g., museums, convention centers). Academic buildings had the highest number of APs, followed by residences, reflecting the university's infrastructure. AP locations are shown in Fig. 1. This categorization informed the analysis of space utilization patterns based on Wi-Fi connectivity data.

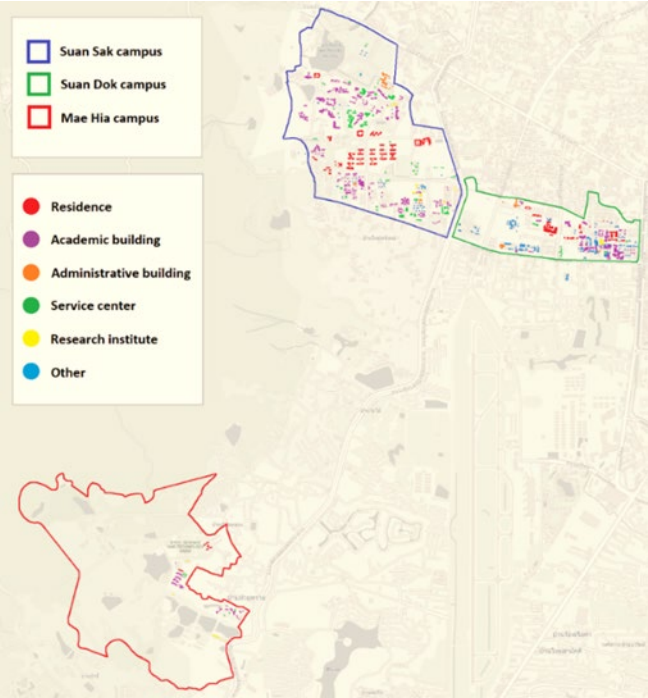


**Figure 1** This study considered the locations of 2,987 Wi-Fi access points across Chiang Mai University's three campuses: Suan Sak, Suan Dok, and Mae Hia

We analyzed campus space utilization through Wi-Fi connectivity data to inform spatial planning and design. Using the Xplaces methodology (Fig. 2), we processed data to extract features such as crowdedness, mobility, and entropy, applying principal component analysis (PCA) and eigendecomposition to reduce dimensionality. X-Means clustering was employed to dynamically determine the number of clusters and their centroids, overcoming the limitations of traditional K-Means by integrating the Bayesian Information Criterion (BIC).

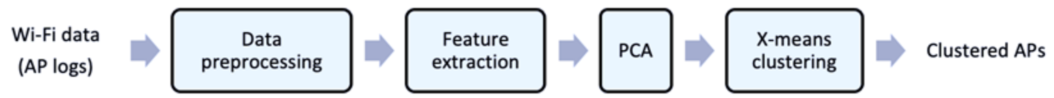


**Figure 2** Summary of our proposed methodology, Xplaces

X-Means iteratively adjusted cluster numbers within a predefined range, retaining the configuration with the best fit. Efficiency improvements were achieved by caching frequently used statistics and focusing on regions requiring adjustments, enabling accurate and dynamic clustering of access points.

Data preprocessing involved removing noise and labeling access points (APs). APs with incorrect or missing locations, such as those situated off-campus, were excluded from the dataset. A custom tool was developed to map AP locations and categorize them into six predefined groups: residence, academic buildings, administrative buildings, service centers, research institutes, and others. The preprocessed dataset for each AP included its ID, geographic coordinates, and connectivity logs, establishing a clean and structured foundation for analysis.

Three key features, crowdedness, mobility, and connectivity entropy, were extracted from the Wi-Fi connectivity data to characterize space utilization. Crowdedness quantified the number

of unique devices connected to an AP within 15-minute intervals. Average crowdedness values were computed across 672 time dimensions, representing quarters of an hour, hours of the day, and days of the week, to capture temporal variations.
Mobility reflected movement dynamics by counting device connections and disconnections during the same 15-minute intervals. Like crowdedness, mobility was aggregated into 672 time dimensions, providing a detailed view of activity patterns.
Connectivity entropy, derived using Shannon's entropy, measured the randomness in connectivity patterns. This metric offered insights into the predictability or complexity of AP usage over time. Entropy was calculated for each AP across the same 672 time dimensions as crowdedness and mobility.
Together, these features provided a comprehensive, multidimensional perspective on campus space utilization, enabling deeper analysis and insights into spatial dynamics.
Using crowdedness, mobility, and connectivity entropy as features, Principal Component Analysis (PCA) was applied to reduce high-dimensional data. PCA projected the features onto principal axes, retaining only the most significant components based on the Scree plot and elbow criterion.
As show in Fig. 3, two principal components were retained for crowdedness and mobility, while three were kept for connectivity entropy and the combined features. The resulting principal components revealed daily patterns: crowdedness and mobility peaked midday, with a secondary rise in the late afternoon, while entropy showed peaks in the morning and afternoon. These components effectively captured the most significant aspects of space utilization dynamics.

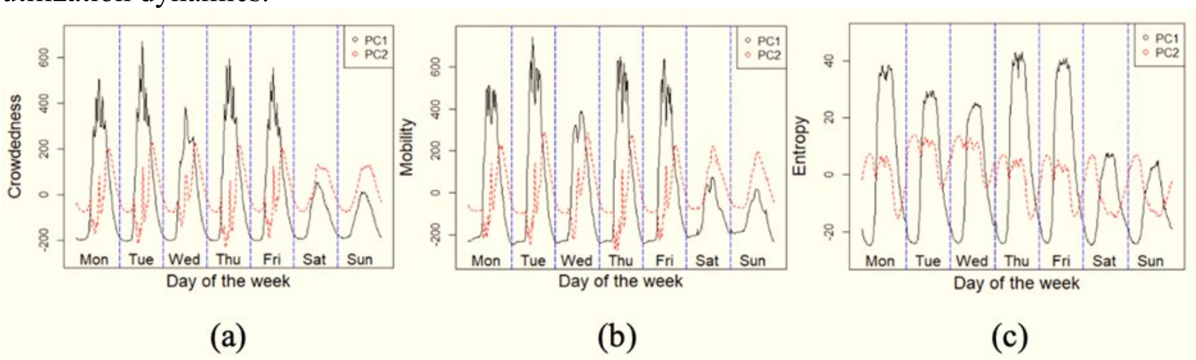


**Figure 3** The primary principal components that capture the (a) crowdedness, (b) mobility, and (c) connectivity entropy features' most significant aspects

We employed eigendecomposition to create a behavioral profile for each access point (AP) based on its connectivity features. By transforming time series data into principal components, we minimized information loss while maximizing variance. Each AP's profile, or signature, was derived from a set of eigenvectors weighted by coefficients, making these signatures comparable and insightful for analyzing space utilization.
To cluster APs, we employed X-means clustering, which overcomes the limitations of traditional k-means by dynamically determining the number of clusters. Guided by the Bayesian Information Criterion (BIC), X-means iteratively splits centroids to optimize cluster formation. Unlike k-means, X-means adjusts clusters locally, updating centroids and redistributing data points until the BIC value is maximized. This approach allowed clusters to emerge naturally based on connectivity patterns, enhancing the analysis of space utilization dynamics.
Extending this exploration further, we investigated the interplay between space utilization and electricity consumption within buildings. This study was conducted using data from two primary sources gathered from 10 academic buildings (totaling an area of 28,224 $m^2$) within

the Faculty of Engineering, Chiang Mai University, Thailand. Both datasets were collected between January 9th and February 3rd, 2020. Electricity consumption data was acquired from energy usage meters, comprising a building ID, energy consumption in kWh, and corresponding timestamps with a 15-minute sampling rate. Wi-Fi network connectivity data originally recorded for network performance monitoring and planning served as a proxy for physical space utilization. This data was collected from 97 Wi-Fi access points (APs) situated within these 10 buildings, also with a 15-minute sampling rate. Each record includes a connected device ID, AP ID, AP geolocation (latitude and longitude), building ID, and its corresponding timestamp.

The locations of the buildings and APs analyzed in this second study are shown in Fig. 4. These buildings serve diverse functions, including lectures, meetings, and research. To account for the influence of room type distribution on building utilization, Table 1 includes the estimated percentages of various room types in each building.

We used the r-squared coefficient to examine the causal relationship between space utilization attributes (crowdedness, mobility, and entropy) and electricity consumption in buildings. For each building, the average values of these attributes across access points were computed and analyzed with time lags ranging from 0 to 60 minutes, in 15-minute intervals.

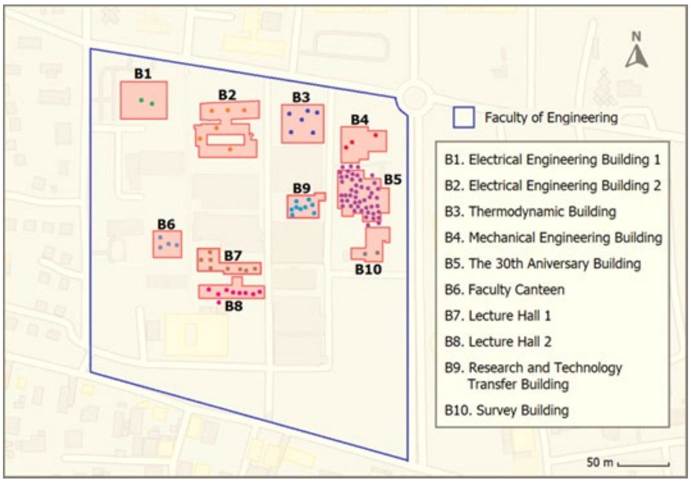


**Figure 4** Locations of the buildings and Wi-Fi access points considered in this study. Access point colors are used to differentiate their belonging to different buildings

**Table 1** The percentage of room types in the building

| **Building** | **Room type** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Lecture** | **Meeting** | **Research lab** | **Student club** | **Academic office** | **Admin office** | **Canteen** |
| B1 | 70 | 0 | 0 | 0 | 0 | 30 | 0 |
| B2 | 40 | 0 | 20 | 10 | 15 | 15 | 0 |
| B3 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| B4 | 0 | 0 | 0 | 0 | 50 | 50 | 0 |
| B5 | 45 | 0 | 25 | 0 | 15 | 10 | 5 |
| B6 | 0 | 0 | 0 | 30 | 0 | 0 | 70 |
| B7 | 60 | 0 | 0 | 0 | 0 | 40 | 0 |
| B8 | 70 | 0 | 0 | 30 | 0 | 0 | 0 |
| B9 | 50 | 50 | 0 | 0 | 0 | 0 | 0 |
| B10 | 60 | 40 | 0 | 0 | 0 | 0 | 0 |

The r-squared values, denoted as $r_u(t-j)$, measured the correlation between electricity consumption and each attribute ($u$) (crowdedness, mobility, or entropy) with a lag of ($j$) minutes. Average r-squared values were calculated for each day of the week to identify the maximum correlations. This process is summarized in Fig. 4.

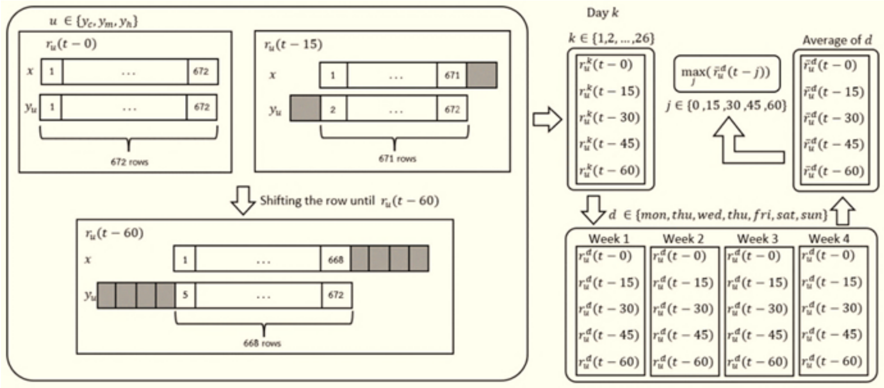


**Figure 5** Overview of the examination of the causal relationship between electricity consumption and space utilization attributes in each building

## RESEARCH RESULTS

We implemented the Xplaces method using four feature sets: crowdedness (C), mobility (M), connectivity entropy (E), and their combination (CME). Results were benchmarked against Calabrese et al.'s eigenplaces method, which uses the number of connections over a 15-minute interval and applies eigendecomposition with k-means clustering.

Clustering effectiveness was evaluated using the Silhouette value, which measures the similarity of data points within their cluster compared to others, ranging from -1 to 1 (higher values indicate better clustering).

Xplaces outperformed eigenplaces, achieving the highest average Silhouette value with E (0.65), followed by M (0.64), CME (0.63), and C (0.53), while eigenplaces scored 0.33, as shown in Fig. 6. These results highlight the superior performance of Xplaces, particularly when using connectivity entropy.

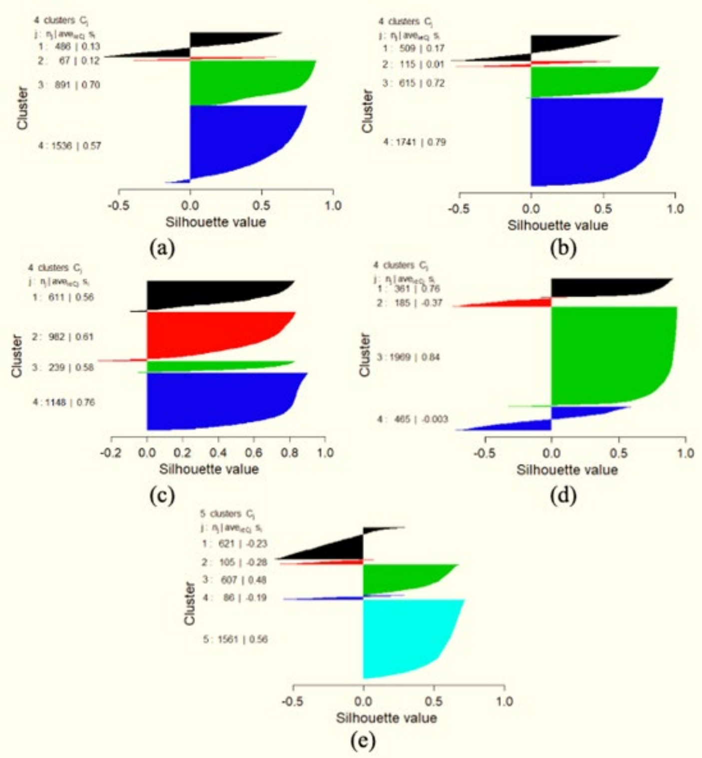


**Figure 6** Silhouette plots depicting the clustering of APs based on Xplaces using (a) crowdedness, (b) mobility, (c) connectivity entropy, (d) combined features, and (e) eigenplaces are displayed

Using connectivity entropy as the feature, Xplaces effectively clustered APs based on their geographical distribution and building types across the three campuses (Figs. 7-9). The method discerned significant patterns in space utilization without reference data, showing that APs in residential areas and academic buildings predominantly clustered together.

The data reflected real-world functionalities of campus spaces, with some areas, such as cafeterias in academic buildings and dormitories, exhibiting similar usage patterns due to shared routines. The clustering closely aligned with real-world activities, particularly distinguishing residential and academic building APs from other types, such as research institutes.


**Figure 7** Clustered APs resulting from Xplaces and their corresponding building types on the Suan Sak campus (main campus)

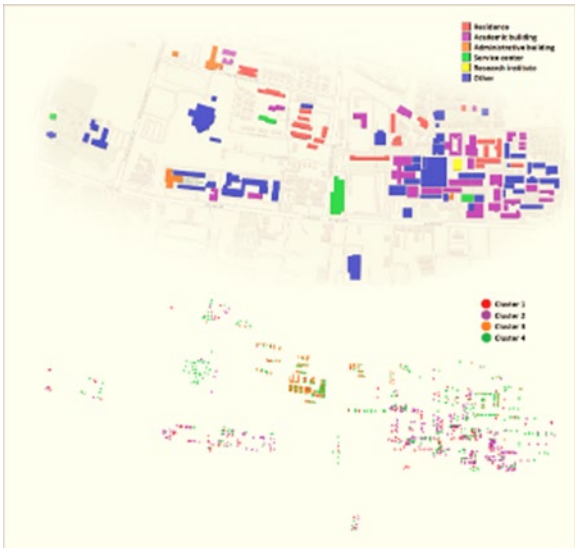

**Figure 8** Clustered APs resulting from Xplaces and their corresponding building types on the Suan Dok campus (health science complex)

**Figure 9** Clustered APs resulting from Xplaces and their corresponding building types on the Mae Hia campus (veterinary medicine and agro-industry faculties)

When looking into how this Wi-Fi based space utilization relates to energy consumption, our analysis of r-squared values across 10 buildings, depicted in Fig. 10, reveals distinct correlations between space utilization and electricity consumption at varying time lags (0-60 minutes) for weekdays and weekends. Correlations were notably higher on weekdays, reflecting regular workday patterns compared to the irregular activities typical of weekends.
For crowdedness, Monday showed the highest correlation at a 45-minute time lag (r-squared = 0.859), indicating that increased crowdedness precedes a rise in electricity consumption by 45 minutes. Similar 30-45 minute timeframes were effective predictors on Tuesday, Wednesday, and Friday, while Thursday demonstrated a shorter 15-minute lag.
Regarding mobility, the strongest correlations occurred at a 15-minute lag on Monday and Thursday, with Wednesday and Friday peaking at 30 minutes. Tuesday mirrored crowdedness with a 45-minute lag.
For entropy, Monday and Friday exhibited the highest correlations at a 30-minute lag, while Tuesday, Wednesday, and Thursday showed a 15-minute lag as the most reliable predictor. These findings highlight the varying timeframes at which space utilization attributes can act as early indicators of electricity consumption, offering valuable insights for energy planning.

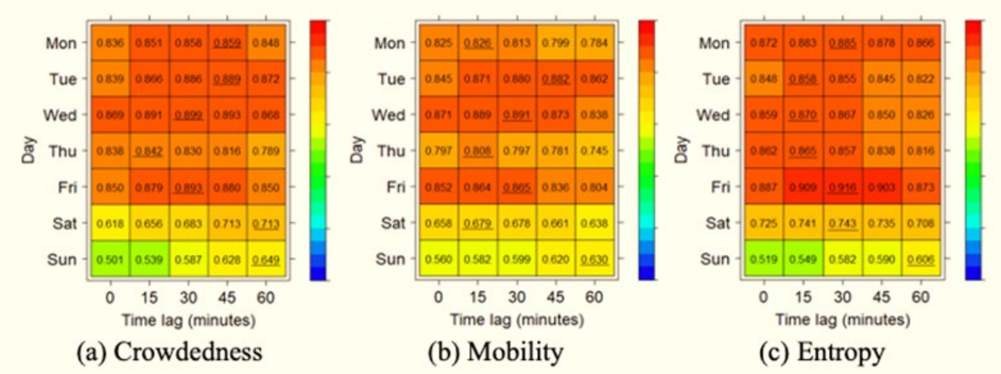


**Figure 10** The heatmap illustrates the r-squared values correlating electricity consumption with (a) crowdedness, (b) mobility, and (c) entropy across various time lags for all buildings combined. The underlined number denotes the highest r-squared value observed for each day of the week across different time lags

# DISCUSSION & CONCLUSION

This study introduces innovative methodologies for analyzing space utilization and its relationship with electricity consumption using Wi-Fi connectivity data. By leveraging features such as crowdedness, mobility, and connectivity entropy, the proposed Xplaces method uncovers nuanced spatial dynamics. Unlike traditional approaches that rely on predefined data structures or extensive reference datasets, Xplaces dynamically clusters access points (APs) and adapts to varying spatial and temporal contexts. The integration of eigendecomposition and X-means clustering further distinguishes this method, enabling flexible, data-driven clustering without requiring prior assumptions about the number of clusters. This adaptability makes Xplaces a versatile tool for urban and building analytics.

The study’s use of opportunistic sensing through Wi-Fi connectivity data is a key innovation. Features like connectivity entropy add a novel dimension to understanding spatial dynamics by capturing the randomness and predictability of occupant behavior. The correlation of space utilization attributes with electricity consumption at varying time lags offers actionable insights for energy management. For instance, identifying the lag between crowdedness and energy spikes enables proactive adjustments to heating, ventilation, air conditioning (HVAC), and lighting systems, reducing energy waste while maintaining occupant comfort. The clustering of APs based on usage patterns also provides valuable insights into space utilization, aiding in the optimization of building layouts and resource allocation.

The implications of this research extend beyond building management to broader urban planning and sustainability efforts. The ability to predict energy consumption patterns based on real-time space utilization data can support dynamic resource allocation and infrastructure design. Additionally, the clustering of spaces based on their usage patterns can inform decisions about space reconfiguration and design improvements, enhancing the functionality and efficiency of urban environments.

The methodologies developed in this study have the potential to transform other domains where understanding human behavior and spatial dynamics is crucial. In healthcare, for example, Wi-Fi data could optimize patient flows and resource allocation, improving operational efficiency and patient care. In transportation, the approach could inform transit planning by analyzing patterns of mobility and crowdedness at transport hubs, supporting schedule optimization and infrastructure development. Similarly, retail and commercial spaces could use this method to analyze customer behavior, enhancing store layouts, inventory management, and marketing strategies. For smart cities, integrating this methodology with other urban sensing technologies, such as IoT devices, could provide a comprehensive view of urban dynamics, enabling informed policy-making and sustainable urban planning.

Future research could explore the integration of additional data sources, such as social media or IoT networks, to enrich the analysis further. Adapting the methodology for predictive modeling through machine learning techniques could anticipate space utilization patterns and their implications for energy consumption. Additionally, addressing privacy concerns through privacy-preserving data analytics would ensure compliance with regulations like GDPR while maintaining the utility of the data.

This study demonstrates the potential of leveraging digital footprints to inform urban and building management, paving the way for smarter, more sustainable spaces. By combining innovative methodologies with practical applications, it offers a robust foundation for future advancements in urban informatics and sustainable development.

## REFERENCES

Araico, A. O. (2017). *Detection of the crowdedness of a place sensing the devices in the area*. M.Sc. thesis, University of Twente.

Calabrese, F., Reades, J., & Ratti, C. (2009). Eigenplaces: segmenting space through digital signatures. *IEEE Pervasive Computing, 9*(1), 78-84.

Cao, X., Dai, X., & Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and buildings, 128*, 198-213.

Cho, H. S., & Choi, M. J. (2014). Effects of compact urban development on air pollution: Empirical evidence from Korea. *Sustainability, 6*(9), 5968-5982.

Erickson, V. L., Carreira-Perpiñán, M. Á., & Cerpa, A. E. (2014). Occupancy modeling and prediction for building energy management. *ACM Transactions on Sensor Networks, 10*(3(42)), 1-28.

Gui, X., Gou, Z., & Zhang, F. (2020). The relationship between energy use and space use of higher educational buildings in subtropical Australia. *Energy and Buildings, 211*, 109799.

Kim, M., & Kotz, D. (2005). Classifying the mobility of users and the popularity of access points. In *Location-and Context-Awareness: First International Workshop, LoCA 2005, Oberpfaffenhofen, Germany, May 12-13, 2005. Proceedings 1* (pp. 198-210). Springer Berlin Heidelberg.

Lemstra, W., Hayes, V., & Groenewegen, J. (2010). *The innovation journey of Wi-Fi: The road to global success*. Cambridge University Press.

Ouf, M. M., Issa, M. H., Azzouz, A., & Sadick, A. M. (2017). Effectiveness of using WiFi technologies to detect and predict building occupancy. *Sustainable Buildings, 2*(7), 1-10.

Sapiezynski, P., Stopczynski, A., Gatej, R., & Lehmann, S. (2015). Tracking human mobility using WiFi signals. *PLoS ONE, 10*(7).

Sevtsuk, A., Huang, S., Calabrese, F., & Ratti, C. (2008). Mapping the MIT campus in real time using WiFi. In *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City* (pp. 326-328).

Shen, L., & Stopher, P. R. (2014). Review of GPS travel survey and GPS data-processing methods. *Transport reviews, 34*(3), 316-334.

Tien, P. W., Wei, S., Calautit, J. K., Darkwa, J., & Wood, C. (2022). Real-time monitoring of occupancy activities and window opening within buildings using an integrated deep learning-based approach for reducing energy demand. *Applied Energy, 308*, 118336.

United Nations. (2014). World Urbanization Prospects. *Demographic Research, 12*(9), 197-236.

Uras, M., Cossu, R., Ferrara, E., Liotta, A., & Atzori, L. (2020). PmA: a real-world system for people mobility monitoring and analysis based on Wi-Fi probes. *Journal of Cleaner Production,* 1-15.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.