

Article

# Geo-Visualization of Spatial Occupancy on Smart Campus Using Wi-Fi Connection Log Data

Zihao Zhao <sup>1,2</sup>, Tao Wang <sup>1,2,\*</sup>, Yiru Zhang <sup>1,2</sup>, Zixiang Wang <sup>1,2</sup> and Ruixuan Geng <sup>1,2</sup>

<sup>1</sup> MOE Lab of 3D Spatial Data Acquisition and Application, Capital Normal University, Beijing 100048, China; zzhao221@cnu.edu.cn (Z.Z.); 2210902166@cnu.edu.cn (Y.Z.); 2220902186@cnu.edu.cn (Z.W.); 2220902129@cnu.edu.cn (R.G.)

<sup>2</sup> College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China

\* Correspondence: wangt@cnu.edu.cn; Tel.: +86-10-68903472

**Abstract:** As a typical and special type of urban setting, the university campus usually faces similar challenges as cities raised by high-density inhabitants. The smart campus has been introduced based on the smart city, as concepts, technologies, and solutions to improve livability and energy efficiency. Inhabitants' occupancy in buildings and open spaces on campus is critical to optimize campus management and services. Information about spatial occupancy of campus inhabitants can be produced based on various location-based solutions, such as global navigation satellite systems (GNSS), campus cameras, Bluetooth, and Wi-Fi. As an essential component in campus information infrastructure, Wi-Fi network covers almost the entire university campus and has advantages in collecting locations of campus inhabitants. In this paper, geo-visualization of spatial occupancy of campus inhabitants is designed and implemented using anonymized Wi-Fi network log data. First, 3-dimension building models are reconstructed based on LiDAR point clouds and construction drawings. Then, the Wi-Fi network log data are cleaned and preprocessed. Campus inhabitants' locations are extracted from structural Wi-Fi data. Geo-visualization at room, floor, and building levels is designed and implemented. On a temporal dimension, spatial occupancy can be visualized by each second, minute, hour, or day of the week in 3D buildings. The implementation of the geo-visualization is based on CesiumJS, which offers an interface for 3D-animated visualization and interaction. The research can be used to support university management and educators to implement the smart campus and optimize pedagogical research.

**Keywords:** geo-visualization; Wi-Fi network; spatial occupancy; smart campus



**Citation:** Zhao, Z.; Wang, T.; Zhang, Y.; Wang, Z.; Geng, R.

Geo-Visualization of Spatial Occupancy on Smart Campus Using Wi-Fi Connection Log Data. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 455.

<https://doi.org/10.3390/ijgi12110455>

Academic Editors: Wolfgang Kainz and Florian Hruby

Received: 26 August 2023

Revised: 31 October 2023

Accepted: 2 November 2023

Published: 6 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Urbanization has been a major driver to advance the economy and society, accompanied with ever emerging global and regional problems, such as traffic, air quality, and noise. The concept of the smart city has been introduced, and relevant technologies have been developed to improve inhabitants' livability and address these challenges in urban areas [1,2]. The university campus is similar in several aspects to cities, which feature a high density of inhabitants and intensive infrastructure usage. Therefore, researchers adopt the concepts, technologies, and solutions of smart cities to implement smart campus initiatives, including the learning environment, energy management, safety and security, and infrastructure management [2]. In the framework of smart cities, geographic information system (GIS) technologies together with geospatial data have been used effectively and geo-visualization is a key interface in nearly all projects of smart cities [3].

Data visualization is an efficient and intuitive way to assist people in interpreting the patterns behind data [4]. Dynamic spatial patterns and geographic knowledge can be uncovered based on the visual exploratory and analytical visualization of large-scale movement data [5]. Trajectory data are a key source for surveillance and management of

mobility agents, whether they are vessels [6] or inhabitant individuals [7]. To implement a successful smart campus, university management needs to understand campus inhabitants' mobility behaviors well in every dimension, in which location and time are two basic factors. Spatial occupancy of campus space is one of the critical indicators, which measures where and when the inhabitants are in each room, floor, building, or open space [8–11]. Dynamic geo-visualization of spatial occupancy can be used to help conduct energy saving [12–15], optimize space utilization [9,10], and analyze risks of epidemic transmission [16,17]. Geo-visualization techniques play an essential and unique role in revealing spatio-temporal patterns of human distribution and utilization of infrastructures [11], making the process of campus management efficient and effective.

In order to calculate spatial occupancy, we basically need to collect people's positions using location-based technologies. A number of approaches for acquiring human positions have been proposed and implemented based on one or a combination of digital instruments, which are either deployed onsite or installed on smart devices. Passive Infrared (PIR) sensors can be mounted at certain locations to monitor and count moving objects in focused areas. It has been well-accepted in relation to privacy issues. However, the extra cost of installation and maintenance [18] and false positive detection [19] block its many applications. Surveillance cameras have been used to derive more detailed information about occupancy pattern [20], but they have a reputation of high intrusiveness into occupants' privacy and require high computational costs [19,21]. Card swiping data collected from security guards or auto-fare collection systems can help count people flow for space occupancy information [22]. However, extra corresponding equipment is necessary, and it is almost impossible to conduct room-level counting in the context of the university campus. Smart devices, such as smart phones and smart watches, equipped with various global navigation satellite systems (GNSS) have been used for acquiring human beings' real-time trajectories. It becomes an effective way of calculating spatial occupancy [23,24]. However, on one hand, it needs to be pre-installed applications on users' devices and on the other hand, is not effective in indoor environments. Research based on call detailed records (CDR) of mobile phone telecommunication data [25,26] can avoid the preinstallation of applications, but its accuracy does not satisfy works on room and building levels.

As a basic information infrastructure service, wireless networks have been available on almost all university campuses in nearly every corner, which support education and research activities. When a mobile device, such as a laptop computer or a smartphone, connects to the wireless network, the networking system records connection information as system log data (Syslog), including the time that this device connects to or disconnects from the network, its Media Access Control (MAC) address, and the name of the access point (AP) with which the device interacts. Basically, each AP has a unique MAC address and is mounted in a specific position. It broadcasts network data over radio frequency (RF) signals and the signal strength varies in space, which can be used to determine the distance between an AP and a smart device connected with the AP. Human position tracking based on a Wi-Fi network has been receiving more and more attention [9,10,14,16,18,19], especially in campus applications.

The occupancy data are usually presented as a set of numeric values, which can be easily visualized by using bar charts [27], line charts [11,22], choropleth maps, graduated symbols on a 2D floor map [11,28], etc. Using a line chart to visualize the occupancy of spatial zones can easily enable a comparison of the occupancy at different times; thereby, facilitating the analysis of changes in occupancy during different periods. However, this approach has certain limitations, particularly when it comes to comparing multiple spatial zones. Geo-visualization of spatial occupancy in different campus spaces can be implemented simply on 2D maps for a holistic overview. However, this approach struggles to capture spatial occupancy of vertical structures in buildings and cannot provide a comprehensive comparison of occupancy across all spatial zones within single or multiple buildings.

In this paper, Wi-Fi Syslog data are used to calculate and generate visualization of spatial occupancy on university campus. First, 3D building models of the campus are

reconstructed based on LiDAR and construction drawings. Then, a Wi-Fi Syslog data preprocessing procedure is designed to extract time-varying online devices. Each AP is pinpointed according to original building CAD (Computer Aided Design) drawings or based on onsite measurement. Each AP is assigned to its corresponding spatial zones. A spatial zone is defined as a room, a floor, a corridor, or a building. Online devices connecting to the Wi-Fi network at each moment of the day change remarkably due to the issues of human behavior day and night, power performance management of the devices, etc. To reduce the uncertainties, a formulated ratio is proposed based on average online devices during the research period. Geo-visualization from the perspective of space, time, and humans is designed and implemented. Based on 3D building models, a 3D geo-visualization system is implemented to generate interactive visualization of spatial occupancy. Our main contribution is to design and implement a 3D geo-visualization method to present spatial occupancy based on structuring Wi-Fi log data.

The rest of the paper is organized as follows. Section 2 introduces the methodology, including data used, reconstruction of 3D building models, preprocessing of Wi-Fi Syslog data, extraction of online devices, and design of the online 3D geo-visualization system. The results are shown in Section 3 for demonstrating the visualization effectiveness of spatial occupancy. Section 4 provides the discussion. Conclusions are presented in Section 5.

## 2. Methodology

### 2.1. Data

The research area is the main campus of Capital Normal University. There are nine teaching and research buildings, eleven dormitory buildings, four administrative buildings, two dining buildings, one library building, and a number of facilities buildings on the campus. The Wi-Fi network covers all campus buildings and nearly all open spaces of the campus. There are over 3000 APs installed in most rooms and open spaces. To reconstruct 3D building models, LiDAR point clouds of the campus and construction drawings of buildings are used.

The Wi-Fi Syslog data span 70 days in a spring semester on one university campus. The data are the output of the network login and logout data in an unstructured text format from the Wi-Fi management system. Basically, each record contains information about the MAC addresses of devices, time of network login and logout, previous AP's name of network connection, current AP's name of network connection, online status of devices, received signal strength indicator (RSSI), operating system (OS) of corresponding devices, connection session length, etc. The MAC address of each device was anonymized by the owner of the data to avoid privacy issues.

Each AP has a unique name on the Wi-Fi network. An AP's name normally indicates the space it is located in. In these cases, an AP can be attached to a specific room or an open space. For those APs that do not bear space information in their names, an AP scanner (Figure 1) based on the ESP8266 microcontroller was built to scan the APs' names, which were then attached to specific spaces.



**Figure 1.** The AP scanner based on ESP8266 microcontroller displaying the MAC address, name, and received signal strength indicator (RSSI) of an AP installed in our lab.

### 2.2. Three-Dimensional Architecture Modelling

The 3D building models of the campus are crucial for the spatialization of devices and the visualization of spatial occupancy. To reconstruct 3D building models, LiDAR point clouds of the campus were first used to determine the envelopes of buildings. The average point density of the LiDAR data is 16 points per square meter. Building footprints were manually extracted. The height of each floor of a building was averaged based on the height of the given building and the floor number of the building. The scanning images of CAD drawings were acquired. The images were georeferenced to align the building footprints from LiDAR point clouds, which are in geographic coordinates. The georeferencing was conducted in QGIS. Each room footprint was digitized semi-automatically into 2D vector data. The height of each room was set to the corresponding floor height. Then, 3D room models were reconstructed based on 2D data and their height. There are a number of rooms or lecture halls crossing multiple floors, and their heights were multiplied by the heights of the floors they crossed. The height of each room is used to conduct an extrusion during visualization. At the same time, each room was attached with attributes, including room name and number, room function, and name list of APs in this room.

The 2D footprints of rooms on the same floor were aggregated to derive the footprints of the given floor. Both rooms and floors can be taken as indoor spatial zones for spatial occupancy visualization.

### 2.3. Wi-Fi Syslog Processing

As indicated above, we may find a number of descriptions of devices connected with the Wi-Fi network. Suppose one connection of a device with Wi-Fi network is considered as an event, which corresponds to one record in the Syslog data. Basically, four types of information can be extracted from the Syslog data: (1) when: the event occurrence time, (2) who: anonymized MAC address of the given device, (3) where: name of the AP(s) the device interacts with, and (4) what: event category which indicates the device joining/leaving the network or roaming from one AP to another AP. Besides the above-mentioned 4Ws, the RSSI values strongly correlate with the distance between devices and APs [29], which can help further determine the position of a given device with a higher accuracy [9]. Devices' operating systems (OS) are also monitored by Wi-Fi network in the Syslog, such as Windows, Android, iOS, or others, which can be used to distinguish the type of a mobile device. Therefore, Syslog related to mobile devices can be extracted based on OS to exclude fixed devices (such as desktop computers, IoT devices, etc.) on occupancy calculation.

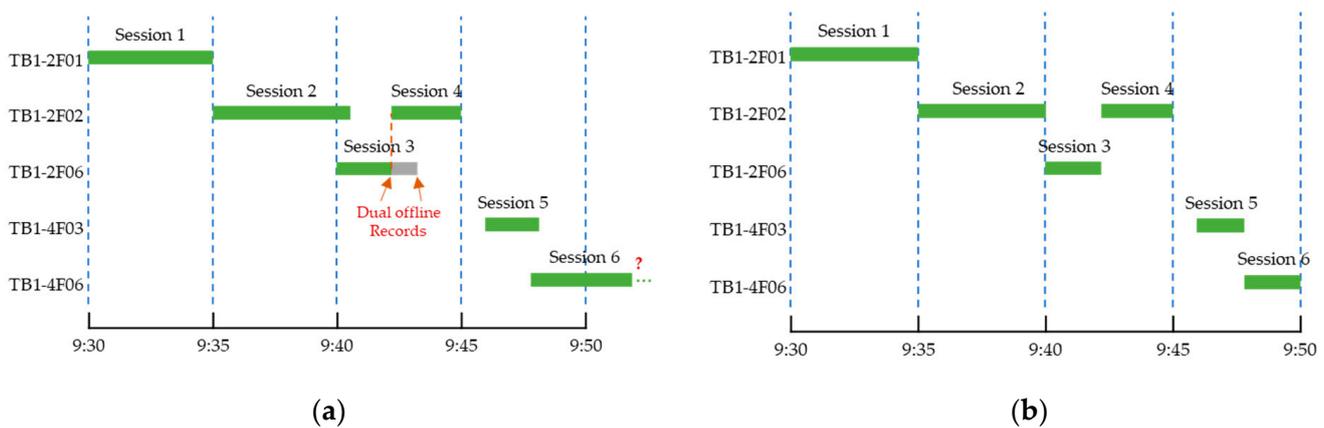
The original Wi-Fi Syslog was provided in an unstructured plain text format. Some of the missing numerical data are filled with various random text placeholders. Therefore, a structural extraction is conducted from the Syslog. The extracting results were exported into a database (Table 1).

**Table 1.** A sample table for one device on a given date extracted from Wi-Fi Syslog data.

Anonymized MAC	Time	Origin AP	Destination AP	Status	RSSI	OS
593e9c2c...	2019-03-01 09:30:00		TB1-2F01	Join	−35	
593e9c2c...	2019-03-01 09:35:00	TB1-2F01	TB1-2F02	Roam	−70	iOS
593e9c2c...	2019-03-01 09:40:00		TB1-2F06	Join	−55	
593e9c2c...	2019-03-01 09:40:32	TB1-2F02		Disconnect	−90	
593e9c2c...	2019-03-01 09:42:24	TB1-2F06	TB1-2F02	Roam	−65	iOS
593e9c2c...	2019-03-01 09:43:14	TB1-2F06		Disconnect	−85	iOS
593e9c2c...	2019-03-01 09:45:00	TB1-2F02		Disconnect	−73	
593e9c2c...	2019-03-01 09:45:44		TB1-4F03	Join	−52	
593e9c2c...	2019-03-01 09:47:40	TB1-4F03	TB1-4F06	Roam	−60	iOS

In order to provide a better description, a session is defined as a time period that a device connects to an AP until it disconnects from this AP. If the device roams from the first AP to another and then logs off from the second one, we define two sessions for this process. For the case in Table 1, each of the three roam records is to be split into two records, an offline record at Origin AP and an online record at Destination AP with the same event time, respectively.

Due to the instability and uncertainty of wireless network connection of mobile devices, there are noises and missing data in Syslog. Time overlaps may occur between neighboring network sessions. Figure 2a shows the status of sessions of the given device, as indicated in Table 1. A shift happened simultaneously from Session 1 to Session 2, which was a reasonable roaming. Session 3 started at 9:40, but Session 2 went beyond 9:40, which means a late report of offline from the AP named TB1-2F02 and Session 2 might end at 9:40. The device roamed from TB1-2F06 to TB1-2F02 at about 9:42 and created an offline record at TB1-2F06. The network system was waiting for the device to send a disassociation message until a timeout happened, and the network forced the device off (again) from TB1-2F06. Therefore, the “dual offline” pattern might happen in the 4W database. Session 4 ended at 9:45, but Session 5 started about 30 s later, which is reasonable due to the re-connection between a device and the network.



**Figure 2.** Illustrative diagram of the synthetic time series data shown in Table 1 (a) with noises and missing data and (b) with cleaned data.

To reduce the uncertainty of data, a data cleaning procedure was designed and implemented to conduct consistency correction among sessions from the table containing 4W information. The pseudo-code of the 4W data cleaning procedure is shown in Algorithm 1 and the cleaned result of Figure 2a is illustrated in Figure 2b.

It happens that some devices might connect to the network before the start time of the Syslog acquisition. Thus, their 4W records might begin with an offline record. Therefore, a makeup of the online record should be performed before data cleaning, and the online time should be set to the starting time of Syslog acquisition. Similarly, the offline record at the very end of the Syslog should also be made up (i.e., Session 6 in Figure 2), but it needs to be performed after the data cleaning procedure because we need to find the exact online record to match it.

#### 2.4. Device Counting and Occupancy Calculation

Campus inhabitants are moving most of the time, which is reflected by devices' moving on network connections. Counting devices connected with APs in a specific room at a specific moment can largely be used to represent the spatial occupancy of inhabitants in this room at this moment. For a specific floor, the results of all rooms and open spaces on this floor can be summed up to derive the corresponding spatial occupancy. Similarly, for a specific building, the results of all rooms and open spaces in this building can be summed up to obtain the building's spatial occupancy at one moment.

**Algorithm 1:** The 4W Data Cleaning Algorithm.**Input:** A list of uncleaned 4W records of a device *RAW4W***Output:** A list of the cleaned 4W records of the device *CLN4W*

```

1  OpenSession = null
2  for each Row in RAW4W do
3      if Row.Status is "Online" then
4          if OpenSession == null then
5              CLN4W.Pushback (Row)
6              OpenSession = Row
7          else
8              makeup an offline record RMakeup of the open session with attributes:
9                  Status = "Offline", Time = Row.Time, AP = OpenSession.AP
10             CLN4W.Pushback (RMakeup)
11             CLN4W.Pushback (Row)
12             OpenSession = Row
13         end if
14     else
15         if OpenSession == null then
16             continue //No corresponding online record exist, consider it as an
17                 exceptional record.
18         else
19             if OpenSession.AP == Row.AP then
20                 CLN4W.Pushback (Row)
21                 OpenSession = null
22             else
23                 continue //Means an offline record of other sessions, consider it
24                     as an exceptional record.
25             end if
26         end if
27     end if
28 end for

```

Generally, the campus population remains stable in daytime and night-time. However, there is a gap in device counting in night-time as some smart devices may be turned off. On another side, larger spaces normally host more people, where more devices can be detected via Wi-Fi network. In this way the exact count of online devices in a specific space cannot reflect its relative spatial occupancy. So, deviation of spatial occupancy from its normalcy is of significance to campus managers. In order to investigate occupancy changes of in a given space to its normalcy, a ratio formula is implemented:

$$R_{dt} = C_{dt} / B_t \quad (1)$$

where  $C_{dt}$  stands for the observed device count at  $t$ th second of day  $d$ ; and  $B_t$  stands for the baseline of normalcy at  $t$ th second, which can best represent the normal condition among the research period. It can be the mean value or the median value of the online device count at the same moment on all dates.  $R_{dt}$  is the ratio of the observed online count to the baseline value, which indicates whether the occupancy is higher or lower than normal.

### 2.5. Geo-Visualization of Spatial Occupancy

Spatial occupancy of different zones (e.g., room, floor, corridor, and building) can be presented in various ways, including statistical tables, bar charts, pie charts, flow charts, and many other types of charts. However, geo-visualization is unique in presenting geospatial patterns and temporal dynamics, which can deliver information more efficiently.

In order to present changes in spatial occupancy at room, floor, corridor, and building levels, geo-visualization at a specific moment and in a specific time period are both important. This study basically uses color schemes to represent the actual value and the density value of spatial occupancy. The values are normalized to 0–1. To enhance the visual saliency of visualizing spatial occupancy in different zones and at different moments, occupancy

data normalization can be implemented with different maximal values (minimal value is 0). The normalization value ( $V$ ) of spatial occupancy can be calculated by:

$$V_{dt} = R_{dt} / \text{Max}(R) \quad (2)$$

where  $R_{dt}$  is the ratio calculated by Formula (1), representing the spatial occupancy of a specific spatial zone; and the  $\text{Max}()$  function calculates a max value for normalization determined by the area of interest or the time of interest. Table 2 presents various conditions for determining maximum occupancy values.

**Table 2.** Procedures for determining maximum occupancy values for normalization under different scenarios.

Compare Between...	One Specific Date	All Dates
Spatial zones of interest	Step 1. Determine which spatial zones are requested to be displayed on the screen; Step 2. Find the maximum occupancy values on the specific date of each spatial zone requested to be displayed; Step 3. Find the maximum occupancy value from the values obtained in Step 2.	Step 1. Determine which spatial zones are requested to be displayed on the screen; Step 2. Find the maximum occupancy values across all dates of each spatial zone requested to be displayed; Step 3. Find the maximum occupancy value from the values obtained in Step 2.
All spatial zones	Step 1. Find the maximum occupancy values on the specific date of all spatial zones in the study area; Step 2. Find the maximum occupancy value from the values obtained in Step 1.	Step 1. Find the maximum occupancy values across all dates of all spatial zones in the study area; Step 2. Find the maximum occupancy value from the values obtained in Step 1.
Individual spatial zone	Step 1. Find the maximum occupancy value on the specific date of the current spatial zone.	Step 1. Find the maximum occupancy value across all dates of the current spatial zone.

A set of colors is attached to a specific spatial zone according to its normalized spatial occupancy at a specific moment. A 3D object of a spatial zone is represented as a 3D polygon with an extruded height and corresponding attributes. A color scheme together with a corresponding timestamp sequence is attached to each 3D object. Original data have a high temporal resolution of up to 1 s. To reduce data volume and network transmission time, a color scheme and its corresponding timestamps can be resampled. A linear interpolation is implemented between two neighboring timestamps to improve visual effect during visualization.

The visualization system was designed as a browser/server structure which enables users to access from computers and portable devices. The programming implementation at the browser side was based on CesiumJS, which is an open-source JavaScript package for 3D visualization [30]. Cesium Language (CZML) is used to accommodate time-sequence 3D geographical data in JSON format and supports streaming data over the internet [31]. A geographical feature is stored as a set of geographic coordinates in WGS84 coordinate system. Time-dynamic color data are stored as a set of RGBA values ranging from 0 to 255. The server side was implemented using C# ASP.NET, which is used to conduct data queries from the database and build the required CZML based on spatio-temporal conditions from user interactions. The server is deployed on a Windows IIS 10.0.

A flow chart illustrating the procedure of 3D geo-visualizing spatial occupancy is shown in Figure 3.

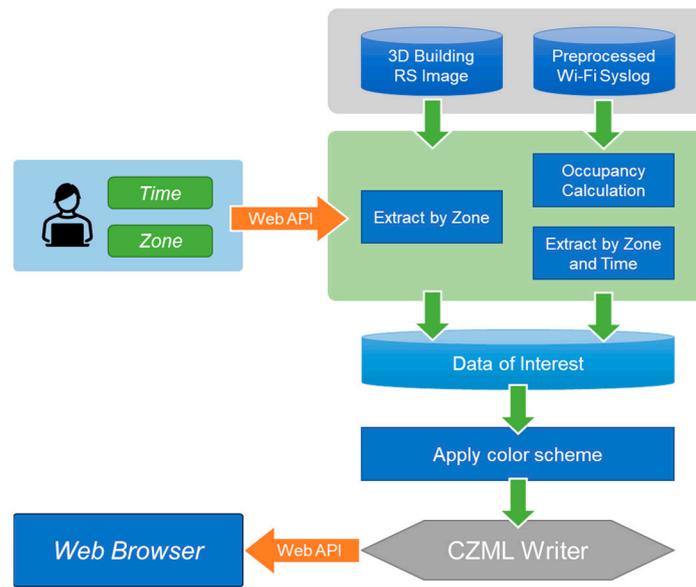


Figure 3. Flow chart of 3D geo-visualizing spatial occupancy.

### 3. Results

Typical statistical charts can present general patterns of human activities. In Figure 4, each curve represents one day for 24 h. The chart shows the overall changing trend of the amount of wireless devices online during the 70-day research period in the research region. Regular patterns can be found in this figure, for example, three prominent peaks in morning, afternoon, and late evening, and two prominent valleys at dinner time and before early morning on workdays.

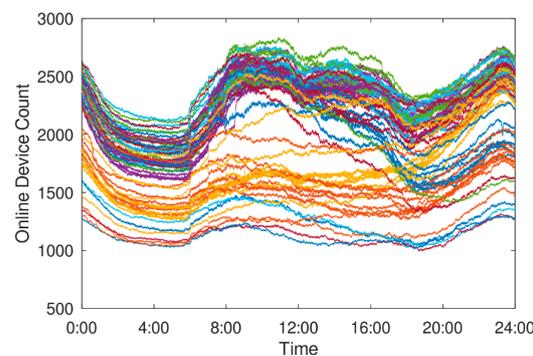
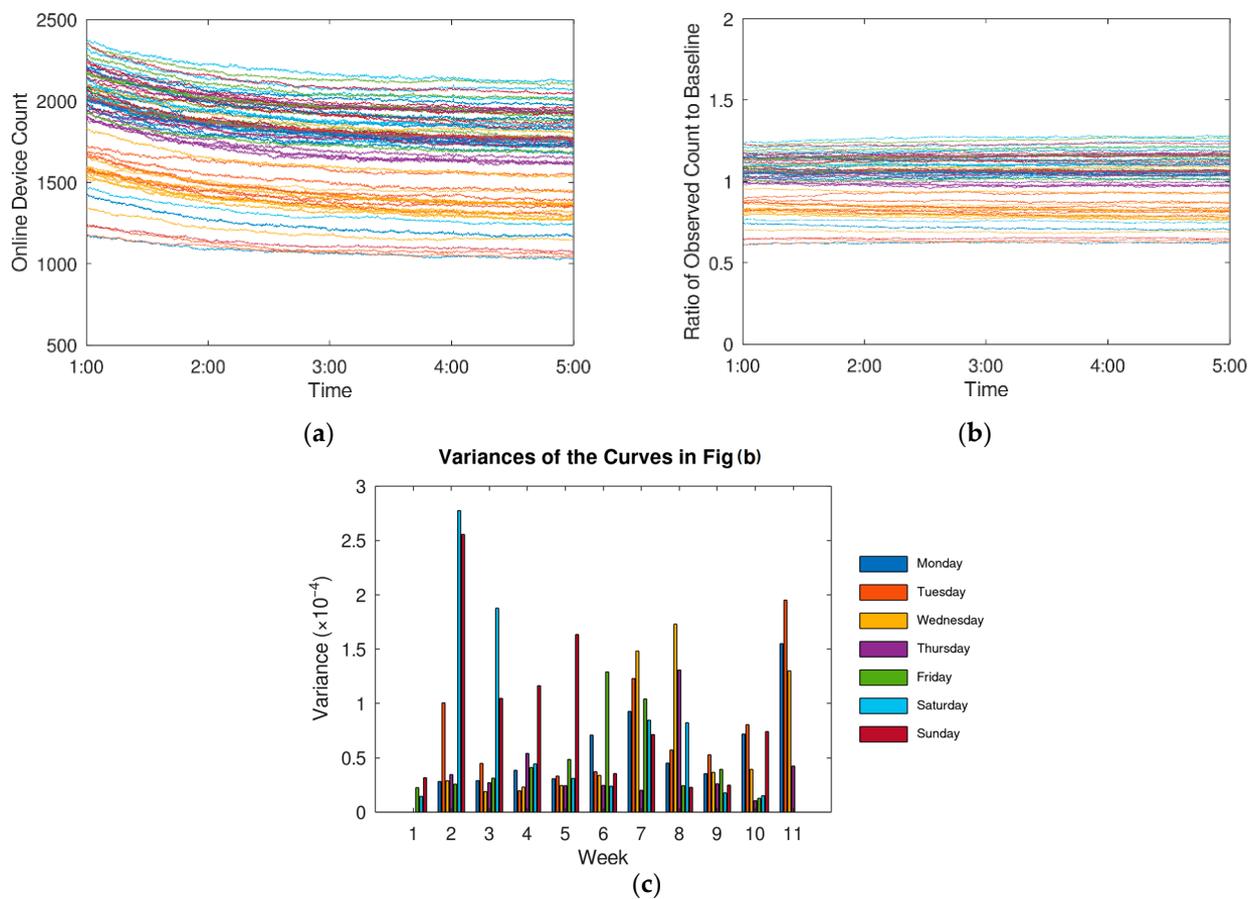


Figure 4. General pattern of the changing total online device count on campus across 70 days.

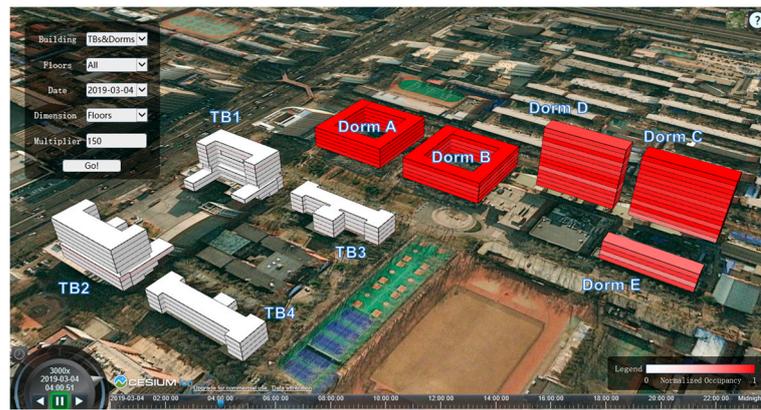
Figure 5a shows the changing trend of the total online device count in campus at night-time (from 1 a.m. to 5 a.m.), which should remain stable. However, the number of devices keeps going down because the devices gradually disconnect from the network. The formula 1 above can be applied using the mean value of the observed counts as a baseline  $B$ , and the result  $R$  is shown in Figure 5b. Then, variances can be calculated to evaluate the smoothness of the curves (Figure 5c). A lower variance indicates a more stable curve, which better reflects the characteristic of night-time fluctuations of campus people.



**Figure 5.** Total online device count on campus at night-time (from 1 a.m. to 5 a.m.) before and after applying the ratio formula. (a) Changing trend of total online device count on campus at night-time; (b) changing trend of the ratio of observed online device count to baseline; (c) variances in the ratio curves evaluating its smoothness.

Major activities on campus are undertaken by students and staff. It is necessary to review typical spatial occupancy patterns of students' dormitory buildings and teaching buildings during work time and night-time, as well as the differences among classrooms with different course schedules. In the research area, there are four teaching buildings (TB1 to TB4). Four dormitory buildings are aggregated to Dorm A and another four to Dorm B for their neighboring positions and internal connectedness. Dorm A and Dorm B host undergraduate students, and the others are for graduate students. Classrooms for the students (primarily undergraduate students) are on the first two floors of TB1 and TB2 and the top floor of TB3 and TB4.

Figure 6 shows the time-varying spatial occupation patterns of building floors at five typical times on a Monday. From Figure 6a, we can find almost all of the crowd was in the dormitory rooms before dawn. Some undergraduate students started their courses at 8 a.m., thus, the floors where classrooms are located are in darker red in Figure 6b. Comparing Figure 6b,c, we can see that some graduate students prefer to start working later than undergraduate students. Figure 6d shows the occupancy pattern after lunch when most undergraduate students return to dormitory buildings because no courses are arranged at noon time. It can also be observed that a part of the graduate students returned to dormitory buildings and had a noon break. Figure 6e is taken at 11 p.m. when teaching buildings stop services. However, a minority of graduate students remained in their labs for research.



(a)



(b)

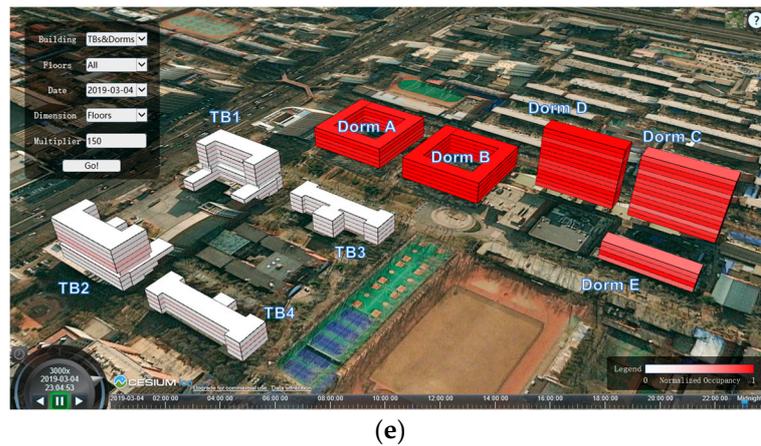


(c)



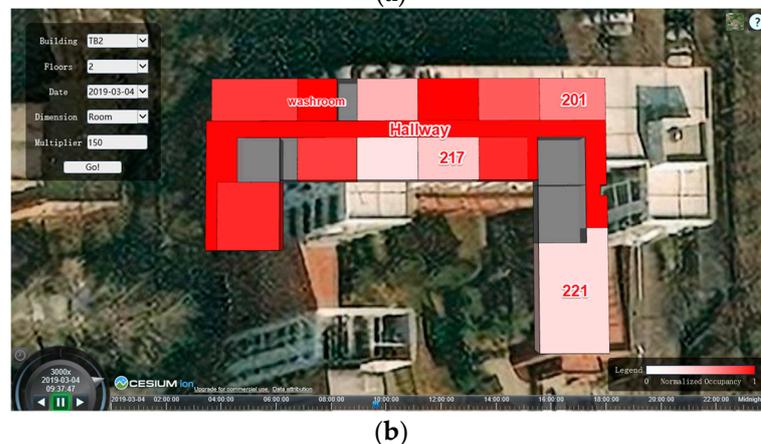
(d)

Figure 6. Cont.



**Figure 6.** Geo-visualization of time-varying occupation patterns on campus: (a) At 4:00 a.m.; (b) at 8:10 a.m.; (c) at 10:40 a.m.; (d) at 12:50 p.m.; (e) at 11:00 p.m.

Figure 7 shows classroom occupancy on the second floor of TB2 on a Monday morning. The course schedule is used as a reference. According to the schedule, rooms 207, 211, and 215 should be unoccupied at 8:40 a.m.(Figure 7a). Room 211 is in dark red, indicating a high occupation beyond the schedule. To evaluate the accuracy of geo-visualization, class interval between 9:30 and 9:40 a.m. can be taken as an example (Figure 7b). According to the schedule, the courses in rooms 201, 217, and 221 should have ended. There is a noticeable change in Rooms 217 and 221. It also shows that the hallway and washroom are in high occupancy. The course in Room 201 ended about 5 min later.



**Figure 7.** Classroom occupancy on the second floor of TB2 on a Monday morning. (a) At 8:40 a.m.; (b) at 9:40 a.m.

#### 4. Discussion

Spatial occupancy on a university campus is important for educators and managers. Previous studies have shown that the device count on the Wi-Fi network can be taken as a proxy for human occupancy estimation and fit with various applications [11]. However, campus inhabitants and visitors may not connect their mobile devices to wireless networks or may carry multiple mobile devices connecting to the wireless network. In addition, the battery management system of smart devices may turn off the Wi-Fi network connection to achieve better battery performance. In some contexts, smart devices in one space may connect APs in a neighboring space due to signal roaming, or even multiple rooms may share one AP [32]. All these bring uncertainties in estimating mobile devices and counting campus population in various 3D spaces. In order to mitigate these challenges, it is necessary to introduce other data sources to improve estimation accuracy. In campus applications, ground truth occupancy data by counting classroom attendance [9,12] or using camera-based occupancy counters [33] can be used. Then, regression models can be built to predict the number of persons in different spatial zones. Although applicable in certain areas on university campuses, these solutions rely on extra sensors or human resources and may not be feasible for a whole university campus, and personal privacy is another issue that should be handled in camera surveillance technologies. It is necessary to acquire accurate positions of smart devices to conduct finer spatial granularity statistics and generate geo-visualization. So RSSI values, Wi-Fi fingerprinting localization, Wi-Fi channel state information, and other techniques can be further introduced. Future works of investigating patterns of crowd behaviors and activities of university campus inhabitants with regular study and work itineraries will be conducted. The ratio formula used to evaluate and reduce the deviation of spatial occupancy from normalcy can be further expanded on an individual or a group of campus inhabitants.

#### 5. Conclusions

Geo-visualization is an effective approach in smart campus, which can assist surveillance, management, and improve service greenness and friendliness on university campus. Spatial occupancy is an essential factor in optimizing infrastructure and ensuring campus security and safety. Geo-visualization of spatial occupancy is more intuitive and efficient than traditional typical statistical charts. To acquire spatial occupancy over university campus, Wi-Fi network Syslog data have advantages compared to data from other location-based methods, which can derive spatial activity information of most campus inhabitants. We designed and developed a procedure for Wi-Fi Syslog data cleaning and structuring for the extraction of spatial occupancy. The 3D geospatial visualization of spatial occupancy in different zones was proposed and implemented.

In this work, we first reconstructed 3D models of campus buildings and then extracted 4W data based on structuring the Wi-Fi Syslog data. A preprocessing procedure was implemented to ensure data consistency and reduce uncertainties. Online devices are used as proxies to calculate the spatial occupancy of buildings at different moments. Geo-visualization of spatial occupancy is designed at room, floor, corridor, and building levels based on normalization schemes in six scenarios. Programming implementation was implemented based on CesiumJS, and CZML was used for online data streaming to support animated 3D geo-visualization. This prototype system can support applications for data analysis, in which campus managers and educators can interpret spatial patterns of individuals in large spaces (e.g., lecture halls or library reading rooms) at various time periods. In data preprocessing, anonymized MAC addresses in the Wi-Fi Syslog data and data aggregation on spatial occupancy calculation reduce privacy risks.

**Author Contributions:** Conceptualization, Zihao Zhao and Tao Wang; methodology, Zihao Zhao and Tao Wang; software, Zihao Zhao and Yiru Zhang; validation, Zihao Zhao, Tao Wang and Zixiang Wang; formal analysis, Zihao Zhao and Tao Wang; investigation, Zihao Zhao and Tao Wang; resources, Tao Wang; data curation, Zihao Zhao, Zixiang Wang and Ruixuan Geng; writing—original draft

preparation, Zihao Zhao and Tao Wang; writing—review and editing, Zihao Zhao and Tao Wang; visualization, Zihao Zhao and Zixiang Wang; supervision, Tao Wang; project administration, Tao Wang; funding acquisition, Tao Wang. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partly funded by the National Natural Science Foundation of China, grant number 41671403.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Bettencourt, L.; West, G. A unified theory of urban living. *Nature* **2010**, *467*, 912–913. [[CrossRef](#)] [[PubMed](#)]
- Verstaavel, N.; Boes, J.; Gleizes, M. From smart campus to smart cities issues of the smart revolution. In Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; pp. 1–6.
- Tao, W. Interdisciplinary urban GIS for smart cities: Advancements and opportunities. *Geo-Spat. Inf. Sci.* **2013**, *16*, 25–34. [[CrossRef](#)]
- Börner, K.; Bueckle, A.; Ginda, M. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1857–1864. [[CrossRef](#)] [[PubMed](#)]
- Andrienko, G.; Andrienko, N.; Dykes, J.; Fabrikant, S.I.; Wachowicz, M. Geovisualization of Dynamics, Movement and Change: Key Issues and Developing Approaches in Visualization Research. *Inf. Vis.* **2008**, *7*, 173–180. [[CrossRef](#)]
- Liu, H.; Chen, X.; Wang, Y.; Zhang, B.; Chen, Y.; Zhao, Y.; Zhou, F. Visualization and visual analysis of vessel trajectory data: A survey. *Vis. Inform.* **2021**, *5*, 1–10. [[CrossRef](#)]
- Nelson, J.K. Cartography & geovisual analytics in personal contexts: Designing for the data creator. *Int. J. Cartogr.* **2023**, *9*, 210–230. [[CrossRef](#)]
- Park, J.Y.; Ouf, M.M.; Gunay, B.; Peng, Y.; O'Brien, W.; Kjærgaard, M.B.; Nagy, Z. A critical review of field implementations of occupant-centric building controls. *Build. Environ.* **2019**, *165*, 106351. [[CrossRef](#)]
- Mohottige, I.P.; Moors, T. Estimating Room Occupancy in a Smart Campus using WiFi Soft Sensors. In Proceedings of the 2018 IEEE 43rd Conference on Local Computer Networks (LCN), Chicago, IL, USA, 1–4 October 2018; pp. 191–199.
- Mohottige, I.P.; Gharakheili, H.H.; Moors, T.; Sivaraman, V. Modeling Classroom Occupancy Using Data of WiFi Infrastructure in a University Campus. *IEEE Sens. J.* **2022**, *22*, 9981–9996. [[CrossRef](#)]
- Oppermann, M.; Munzner, T. Ocupado: Visualizing Location-Based Counts Over Time Across Buildings. *Comput. Graph. Forum* **2020**, *39*, 127–138. [[CrossRef](#)]
- Ashouri, A.; Newsham, G.R.; Shi, Z.; Gunay, H.B. Day-ahead Prediction of Building Occupancy using WiFi Signals. In Proceedings of the 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, BC, Canada, 22–26 August 2019; pp. 1237–1242.
- Soetedjo, A.; Sotyohadi, S. Modeling of Occupancy-Based Energy Consumption in a Campus Building Using Embedded Devices and IoT Technology. *Electronics* **2021**, *10*, 2307. [[CrossRef](#)]
- Jagadeesh Simma, K.C.; Mammoli, A.; Bogus, S.M. Real-Time Occupancy Estimation Using WiFi Network to Optimize HVAC Operation. *Procedia Comput. Sci.* **2019**, *155*, 495–502. [[CrossRef](#)]
- Balaji, B.; Xu, J.; Nwokafor, A.; Gupta, R.; Agarwal, Y. Sentinel: Occupancy Based HVAC Actuation Using Existing WiFi Infrastructure within Commercial Buildings. In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, Roma, Italy, 11–15 November 2013.
- Zakaria, C.; Trivedi, A.; Cecchet, E.; Chee, M.; Shenoy, P.; Balan, R. Analyzing the Impact of COVID-19 Control Policies on Campus Occupancy and Mobility via WiFi Sensing. *Acm Trans. Spat. Algoritm. Syst.* **2022**, *8*, 22. [[CrossRef](#)]
- Li, J.; Sharma, A.; Mishra, D.; Batista, G.; Seneviratne, A. COVID-Safe Spatial Occupancy Monitoring Using OFDM-Based Features and Passive WiFi Samples. *Acm Trans. Manag. Inf. Syst.* **2021**, *12*, 24. [[CrossRef](#)]
- Melfi, R.; Rosenblum, B.; Nordman, B.; Christensen, K. Measuring building occupancy using existing network infrastructure. In Proceedings of the 2011 International Green Computing Conference and Workshops, Orlando, FL, USA, 25–28 July 2011; pp. 1–8.
- Zou, H.; Zhou, Y.; Yang, J.; Spanos, C.J. Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT. *Energy Build.* **2018**, *174*, 309–322. [[CrossRef](#)]
- Sun, K.; Zhao, Q.; Zhang, Z.; Hu, X. Indoor occupancy measurement by the fusion of motion detection and static estimation. *Energy Build.* **2022**, *254*, 111593. [[CrossRef](#)]
- Zou, H.; Jiang, H.; Yang, J.; Xie, L.; Spanos, C. Non-intrusive occupancy sensing in commercial buildings. *Energy Build.* **2017**, *154*, 633–643. [[CrossRef](#)]
- Wei, L.; Guo, D.; Chen, Z.; Yang, J.; Feng, T. Forecasting Short-Term Passenger Flow of Subway Stations Based on the Temporal Pattern Attention Mechanism and the Long Short-Term Memory Network. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 25. [[CrossRef](#)]

23. Kim, S.; Lee, E.; Park, S.; Kwon, K. Non-intrusive Occupancy Estimation using a Smartphone GPS. In Proceedings of the 2021 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, 20–22 October 2021; pp. 1853–1855.
24. Shi, H.; Huang, H.; Ma, D.; Chen, L.; Zhao, M. Capturing urban recreational hotspots from GPS data: A new framework in the lens of spatial heterogeneity. *Comput. Environ. Urban Syst.* **2023**, *103*, 101972. [[CrossRef](#)]
25. Ghahramani, M.; Zhou, M.; Hon, C.T. Mobile Phone Data Analysis: A Spatial Exploration Toward Hotspot Detection. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 351–362. [[CrossRef](#)]
26. Maji, G.; Mandal, S.; Sen, S. Identification of City Hotspots by Analyzing Telecom Call Detail Records Using Complex Network Modeling. *Expert Syst. Appl.* **2023**, *215*, 119298. [[CrossRef](#)]
27. Verbree, E.; van der Spek, S.; Kalogianni, E.; Ailerytè, R.; Lam, M.; Zhou, K.; van der Ham, M. Passive WiFi Monitoring of the Rhythm of the campus. In Proceedings of the 18th AGILE International Conference on Geographic Information Science, Lisboa, Portugal, 9–12 June 2015.
28. Min, S.; Ahuja, R.; Liu, Y.; Zaidi, A.; Phu, C.; Nocera, L.; Shahabi, C. CrowdMap: Spatiotemporal Visualization of Anonymous Occupancy Data for Pandemic Response. In Proceedings of the 29th International Conference on Advances in Geographic Information Systems, Beijing, China, 2–5 November 2021; pp. 630–633.
29. Ravi, A.; Misra, A. Practical server-side WiFi-based indoor localization: Addressing cardinality & outlier challenges for improved occupancy estimation. *Ad Hoc Netw.* **2021**, *115*, 102443. [[CrossRef](#)]
30. Cesium GS, Inc. Why Cesium. Available online: <https://cesium.com/why-cesium/> (accessed on 23 February 2023).
31. Hunter, S. Cesium Language (CZML) Guide. Available online: <https://github.com/AnalyticalGraphicsInc/czml-writer/wiki/CZML-Guide> (accessed on 23 February 2023).
32. Ren, Y.; Salim, F.D.; Tomko, M.; Bai, Y.B.; Chan, J.; Qin, K.K.; Sanderson, M. D-Log: A WiFi Log-based differential scheme for enhanced indoor localization with single RSSI source and infrequent sampling rate. *Pervasive Mob. Comput.* **2017**, *37*, 94–114. [[CrossRef](#)]
33. Alishahi, N.; Ouf, M.M.; Nik-Bakht, M. Using WiFi connection counts and camera-based occupancy counts to estimate and predict building occupancy. *Energy Build.* **2022**, *257*, 111759. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.