A Combining Model for Crowd Flow Prediction in Partition Coverage Wireless Network Scenarios

Zheng Ma[1], Rui Zhang[2], Lang Gao[3,*]

{mazheng@cug.edu.cn[1], zhangrui@cug.edu.cn[2], langgao@cug.edu.cn[3]}

Informatization Office, China University of Geosciences, Wuhan, China[1]
Informatization Office, China University of Geosciences, Wuhan, China [2]
Human Resources Department, China University of Geosciences, Wuhan, China[3]
*corresponding author

Abstract. With the development of modern higher education, university campuses serve multiple functions for both faculty and students, including teaching, research, and daily life. These campuses commonly face issues such as high population density, frequent mobility, and complex composition. Accurate prediction of campus traffic environments is fundamental for scientific and efficient management. Predictive data can support various aspects of campus operations, including energy conservation, resource allocation, event scheduling, hazard prevention, and emergency response, making it of significant importance. This paper leverages the advantages of wireless network seamless roaming technology to obtain the number of people in different areas from wireless controllers. Based on this, we propose a crowd flow prediction model using ARIMA-CNN-LSTM. The model has been tested, and its accuracy meets expectations and surpasses traditional algorithms.

Keywords: wireless network, crowd prediction, ARIMA, CNN-LSTM.

1 Introduction

Universities serve as cradles for nurturing high-quality talents and are crucial bases for knowledge innovation and scientific research. Campus management is a vital component in ensuring the normal operation of universities and achieving their educational objectives. The university campus, which hosts tens of thousands of faculties and students for learning, research, work, and daily life, typically has a high population density. Accurate monitoring of personnel movement within the campus is essential for enhancing the level of campus management.

Predicting pedestrian flow in campus public areas can help students plan and arrange their time in advance, avoiding peak hours when visiting libraries, dining halls, and other areas, thereby ensuring the effective utilization of public resources. By forecasting the number of diners, it aids staff in making reasonable scheduling and arrangements, reducing the potential risks associated with large gatherings and providing staggered services. By considering pedestrian trajectories, it is possible to better predict the air conditioning load of building clusters, effectively reducing waste in building air conditioning energy consumption. In summary, predicting the flow of people on campus holds significant importance for enhancing campus safety management, optimizing resource allocation, achieving energy conservation and emission reduction, and improving emergency response capabilities.

CONF-MLA 2024, November 21, Adana, Turkey

DOI 10.4108/eai.21-11-2024.2354639

In recent years, with the continuous increase in investment in the construction of information infrastructure in universities, campuses have generally achieved full coverage of wireless network signals. Wireless networks, with their flexibility and portability, provide users with the ability to access the internet anytime and anywhere. How can wireless network technology be utilized to assist in predicting campus crowd flow? A significant amount of research has been conducted in the field of crowd flow prediction, resulting in the development of various prediction models and methods.

In this paper, we propose a regional crowd flow prediction model that integrates WiFi technology. This model predicts regional crowd flow based on the number of users connected to the wireless network in a given area, offering a new solution for predicting campus personnel flow. The structure is organized as follows: Section 2 introduces the research work and literature review in the relevant field. Section 3 presents the correlation between the deployment structure of campus wireless networks and regional crowd flow. Section 4 introduces algorithm principles and establishes a crowd flow combination prediction model. Section 5 presents and evaluates the experimental results. Finally, Section 6 concludes the paper and points out further work.

2 Related work

A variety of techniques-based approach for crowd flow prediction has been proposed. Z. Chen et al. [1] presented a passive human activity recognition approach using WiFi CSI data processed by an Attention-Based BLSTM network, enabling activity detection without the need for wearable sensors. P. Fuxjaeger et al. [2] discussed a method for privacy-preserving Wi-Fi monitoring to analyze road traffic without compromising the privacy of individual devices, leveraging the high number of probe requests emitted by vehicles for meaningful travel time information. M.Traunmueller et al. [3] presented a model that uses large-scale WiFi probe request data to simulate urban mobility trajectories in dense urban environments, highlighting the potential of WiFi data for understanding city mobility patterns and the privacy issues associated with the growing availability of public WiFi networks. L. W. Chen et al. [4] developed a WiFi-based monitoring system to analyze campus crowd behavior by examining MAC addresses, RSSI, and Wireless APs, employing time series analysis for predictions and offering insights into crowd dynamics. C. H. Liu et al. [5] presented an Attentive Convolutional LSTM model for citywide crowd flow modeling, which integrates spatial and temporal features to predict crowd density and flows across different locations in a metropolitan area. M. V. Ramesh et al. [6] introduced a novel wireless sensor network architecture that leverages smartphones as sensor nodes to predict and mitigate crowd disasters, particularly stampedes, by analyzing data from embedded sensors like accelerometers, GPS, and light sensors. O. Yamada et al. [7] presented a crowd flow prediction method that utilizes time series Point of Interest (PoI) stay counts, aiming to support behavior planning for crowd avoidance, and demonstrates its effectiveness in reducing prediction errors compared to methods based solely on headcount. K. Yasufuku and A. Takahashi [8] have developed a real-time crowd flow prediction and visualization platform that integrates agent-based crowd simulation with advanced crowd management systems, designed for crowd management in large events and specific facilities. Xing, J. et al. [9] propose a prediction method based on Graph Neural Networks (GNN), referred to as STGs (Spatial-Temporal Graphs), for crowd flow prediction within urban areas. This approach involves jointly constructing spatial and temporal graphs from grid maps, and then applying Graph Neural Networks to directly capture the relationships between regions. D. He et al. [10] proposed a spatiotemporal hybrid neural network model for crowd flow prediction in critical urban areas, termed NDV-LSTM. This model integrates the Node2Vec graph embedding algorithm with the Long Short-Term Memory (LSTM) network to predict crowd flow in key urban regions. H. Go et al. [11] proposed a deep learning model based on global-local structures for crowd flow prediction. This model simultaneously utilizes both overall (global) and location-type-specific (local) crowd flow data to enhance the prediction accuracy for each sub-group. The research results indicate that the proposed model improves prediction accuracy across different datasets compared to related work that only uses target category data for prediction. Shiyu Zhang et al. [12] proposed a framework called CrowdTelescope, which is based on Wi-Fi positioning data for multi-granularity spatiotemporal crowd flow prediction in smart campuses. They designed a robust method to extract human mobility trajectories from noisy Wi-Fi connection records and employed spatiotemporal graph neural networks to model crowd flow under different granularity levels of location hierarchies. Additionally, they developed a prototype system that visualizes historical and predicted crowd flows within the campus through an interactive map. The study results demonstrate that CrowdTelescope can effectively extract information-rich human mobility trajectories from noisy Wi-Fi connection records and accurately predict crowd flows at various granularity levels across the campus. C. W. Hu et al. [13] proposed a dynamic graph convolutional network model named Res-DGCN for predicting human mobility dynamics in urban areas. The model captures multi-order spatial relationships through a spatial attention module, learns temporal trends through a conditional convolution module, and employs a Huber loss function to reduce sensitivity to outliers. On two public datasets, Res-DGCN achieved significant improvements in prediction accuracy compared to baseline models. M. Luca et al. [14] have provided a classification of the applications of deep learning in predicting and generating human mobility tasks, including next location prediction, crowd flow prediction, trajectory generation, and flow generation. The challenges associated with each task are discussed, as well as how deep learning overcomes the limitations of traditional models, offering the most relevant solutions. B. F. Yao et al. [15] introduce a machine learning-based approach for constructing a pedestrian tracking system in smart cities. The article employs the YOLOv3 algorithm for pedestrian detection, combined with the Mean Shift algorithm and the Kalman filter for tracking, to count the number of pedestrians entering and leaving specific areas. Experimental results demonstrate that the system can effectively detect and track pedestrians, which is of significant importance for pedestrian flow statistics in public areas of smart cities. Q. S. Teng et al. [16] proposed a novel deep learning model named STA-CFPNet for predicting pedestrian flow in indoor areas. The model consists of four branches, each handling temporal proximity, periodicity, trend, and external factors, respectively. By fusing the outputs of these branches, the model aims to uncover temporal dependencies within the data. Additionally, the study introduced a spatial-temporal attention-based building block to enhance the model's convergence speed and prediction performance.

Current research has extensively focused on crowd flow prediction in the fields of smart cities and smart transportation, with various high-precision prediction models proposed. However, studies specifically targeting crowd flow prediction within university campuses are relatively scarce. The deployment of wireless networks in university campuses provides a foundational environment for predicting crowd movements based on wireless coverage areas. Predictive research using wireless network data can offer insights for smart campus management, thereby holding significant research value.

3 Campus Wireless Network

3.1 Partition Coverage Wireless Network

The Wireless Campus Network is a system that provides wireless internet access services within the campus, allowing faculty and students to connect to the internet from anywhere on campus using wireless devices such as laptops, smartphones, and tablets. Below are some fundamental principles and components of the structure of a wireless campus network:

Wireless Access Points (AP). Wireless Access Points are the foundation of the wireless network. They are responsible for transmitting wireless signals, enabling wireless devices to connect to the network. Each access point has a coverage area, commonly referred to as a "hotspot," within which devices can connect to the network.

Wireless Access Controllers (AC). In large-scale networks, wireless controllers are used to centrally manage multiple wireless access points, providing unified configuration, monitoring, and management.

Core Switches. Core switches serve as the central hub of the network, handling large volumes of data transmission to ensure high-speed and stable network performance.

Authentication Systems. Campus wireless networks typically require authentication to ensure that only authorized users can access the network. This may include username and password, digital certificates, multi-factor authentication, and other methods.

As shown in Figure 1, in large campuses, the campus wireless network can be divided into different wireless zones based on actual needs and geographical layout, such as residential areas, libraries, and teaching areas. Each zone deploys an independent wireless controller, which is configured according to its specific requirements and environmental characteristics.

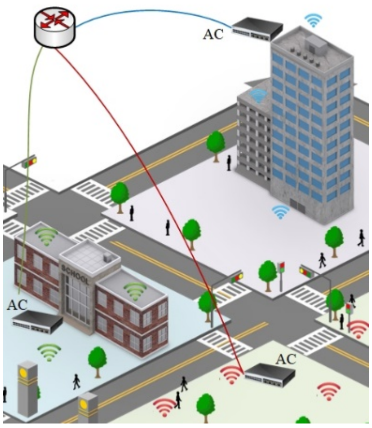

Fig. 1. Diagram of wireless network structure

3.2 User Terminal Roaming Principle

To further enhance the online experience for faculty and students, the concept of seamless authentication for internet access has been proposed and improved upon, building on existing methods such as AAA client authentication, Portal authentication, and 802.1x authentication. This approach is widely adopted by universities as a common method for internet authentication. Seamless authentication for internet access is a technology that allows users to authenticate almost imperceptibly, primarily by recording the MAC address of the user's terminal as a subsequent authentication identifier, thereby enabling rapid authentication for campus networks. This method means that once a user's wireless device successfully logs into the campus network with seamless authentication enabled for the first time, the system automatically records the device's MAC address information. Subsequently, when the wireless device attempts to access the campus network again or moves across different areas, no further login operations are required.

As illustrated in Figure 2, in a campus environment equipped with coverage-unaware authentication wireless networks, user terminal devices can automatically associate with access points to maintain continuous network access. This technology not only enhances user experience but also provides the foundational conditions for collecting personnel mobility data. Given that people commonly carry smartphones with them, analyzing the data from wireless controllers across different areas allows for a certain degree of crowd flow prediction.

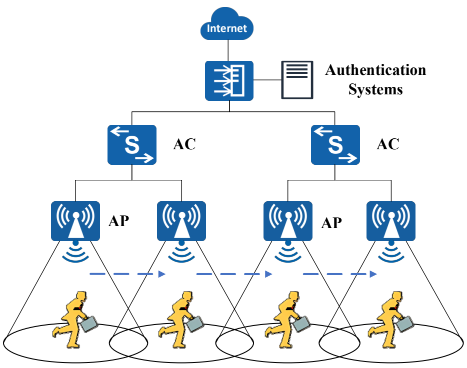

Fig. 2. Maintaining wireless network connectivity for mobile users

4 Crowd Flow Prediction Model

4.1 ARIMA and CNN-LSTM

The ARMA model, or Autoregressive Moving Average model, is a key method for analyzing time series. It is a "hybrid" model, based on both the autoregressive (AR) model and the moving average (MA) model. The formula can be expressed as:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{p-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

Differentiation is a preprocessing technique utilized to render non-stationary time series stationary. Taking the first difference in a time series involves comparing each observation with the one preceding it and calculating the difference between them. For instance, given a time series Y_t, the first difference of the series can be defined as the difference between Y_t and Y_{t-1}.

$$\Delta Y_t = Y_t - Y_{t-1} \quad (2)$$

The ARIMA model is suitable for non-stationary time series. After differencing to achieve stationarity, it utilizes autoregressive and moving average terms to capture the dynamic characteristics and random fluctuations of the series. This model is widely applied in predictive analysis within the fields of economics and finance.

The CNN-LSTM model is a deep learning architecture that synergizes the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) for processing data with both spatial and temporal characteristics. The CNN component is primarily employed to extract spatial features from the data, while the LSTM component is utilized to handle time series data, capturing temporal dependencies. This combined model has demonstrated exceptional performance across various domains, including video analysis, time series forecasting, and natural language processing.

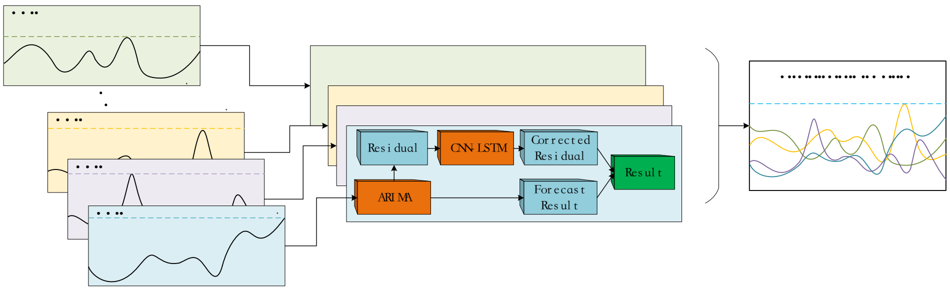

Fig. 3. The model for crowd flow prediction

4.2 The Combining Model for Crowd Flow Prediction

To predict future campus crowed flow, we establish the model based on ARIMA-CNN-LSTM. The prediction process of the model is illustrated in Figure 3.

1. Data Collection: User information is periodically collected from each wireless controller through the SNMP protocol or network management platform.

2. Data Normalization: Information from wireless controllers of different brands and models may require standardized preprocessing.

3. Preprocessing with ARIMA: Apply the ARIMA model to preprocess the raw crowd flow time-series data. This step aims to remove non-stationarity from the time series, converting it into a stationary sequence.

4. Residual Extraction: Extract the residual values between the raw data and the ARIMA prediction results to obtain the temporal dependencies and trends in the residuals.

5. Feature Extraction and Prediction with CNN-LSTM: Based on the residual data over n time units, the CNN-LSTM model performs feature extraction and prediction, yielding the corrected residual data.

6. Correcting results: The ARIMA prediction results with corrected residual data to form the final time-series prediction results.

7. Generate prediction results: Combine the prediction results of each wireless controller to form a complete regional population flow prediction result.

5 Experiment and Evaluation

5.1 Dataset and Environment

The dataset is compiled from historical data of a real campus network. The data pertains to the wireless network operation of an independent campus over a one-month period. As shown in Table I, the campus wireless network is divided into five zones, each with a distinct grouping of wireless access points (APs) for coverage. The collection period includes weekdays, weekends, and holidays, making the data representative of typical scenarios.

Table 1. Wireless network coverage area division

area id	*ac id*	*description*
a	ac1	public teaching buildings
b	ac2	laboratories and office buildings
c	ac3	block 1 student apartment complex
d	ac4	block 2 student apartment complex
e	ac5	campus roads and outdoor public areas

5.2 Model Evaluation

To reflect the degree of approximation between predicted values and actual values, this paper employs the Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE) to evaluate whether the prediction model can improve prediction accuracy. The formulas for R^2, and MAPE are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \widehat{Y_i})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \tag{3}$$

$$E_{MAP} = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - \widehat{Y_i}|}{Y_i} \times 100\% \tag{4}$$

where: n is the number of samples, Y_i is the measured crowd flow; $\widehat{Y_i}$ is the predicted crowd flow, Y is the mean of the actual crowd flow. Smaller values of MAPE, or R^2 closer to 1, indicate better model prediction performance.

5.3 Experimental Test Results

The proposed temporal prediction model was tested using a real dataset. The original data, the results of ARIMA model processing, and the residuals obtained from the ARIMA model are shown in Figure 4-8. The residuals fluctuate around 0, indicating that the model effectively captures the dynamic characteristics of the data.

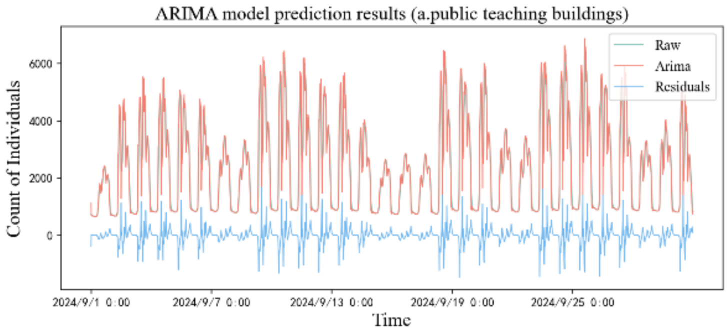

Fig. 4. ARIMA prediction and residuals(a)

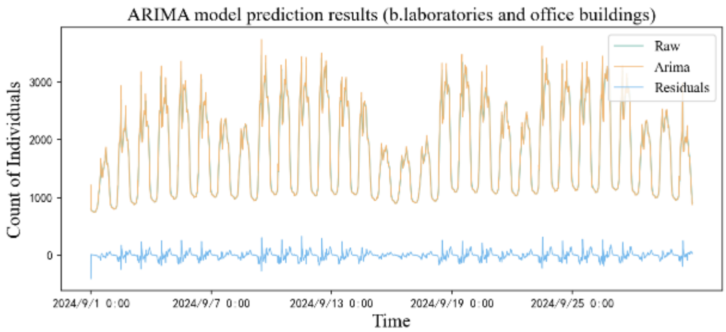

Fig. 5. ARIMA prediction and residuals(b)

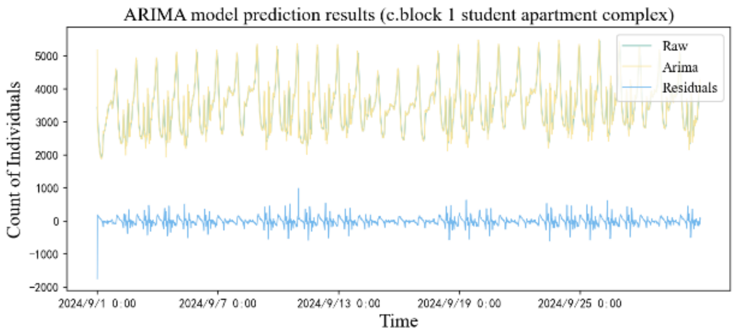

Fig. 6. ARIMA prediction and residuals(c)

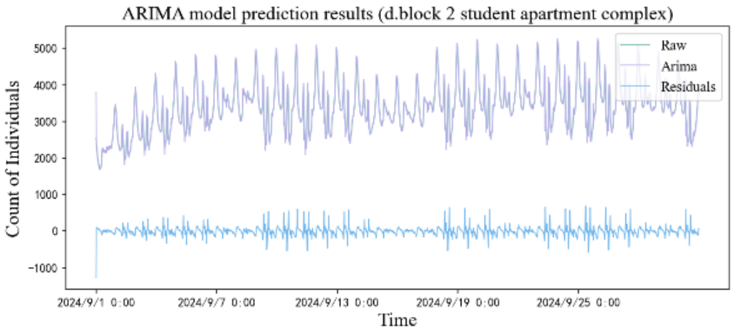

Fig. 7. ARIMA prediction and residuals(d)

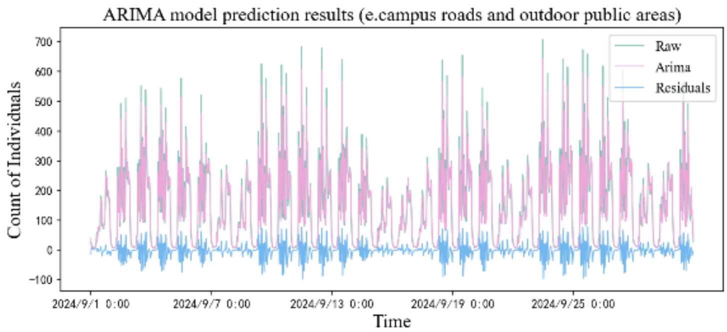

Fig. 8. ARIMA prediction and residuals(e)

Figure 9 presents the combined prediction results. Compared to other models, we observe that the ARIMA-CNN-LSTM model exhibits the highest R2 value. Table II details the experimental results.

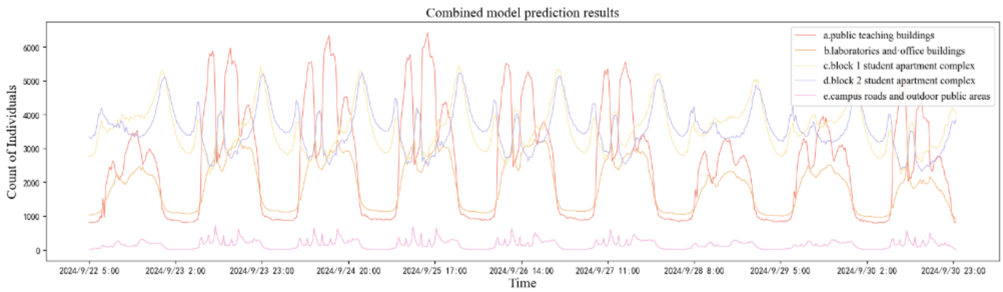

Fig. 9. Combined model prediction results

Table 2. Comparison of time model performance

model	*a*		*b*		*c*		*d*		*e*	
	R^2	E_{MAP}	R^2	E_{MAP}	R^2	E_{MAP}	R^2	E_{MAP}	R^2	E_{MAP}
CNN	0.982	8.59%	0.995	1.73%	0.985	1.66%	0.981	1.73%	0.893	2.86%
LSTM	0.986	4.73%	0.985	2.51%	0.976	1.96%	0.959	2.13%	0.860	4.15%
CNN-LSTM	0.994	4.11%	0.990	2.54%	0.971	2.47%	0.978	2.18%	0.934	2.47%
ARIMA-CNN-LSTM	0.996	2.29%	0.999	0.75%	0.994	1.04%	0.994	0.96%	0.997	6.6%

5.4 Discussion

Through real-data validation, the crowd flow model based on ARIMA-CNN-LSTM achieves highly accurate predictions. This model is applicable to campus network zoning scenarios and demonstrates high prediction accuracy across different wireless controller coverage areas, with R^2 values exceeding 0.994 for crowd predictions in five regions. Comparative experiments with other single models also show that this combined algorithm performs superiorly.

6 Conclusions

The proposed combined model for predicting crowd flow in campus areas based on wireless network data, after being tested with real datasets, has achieved the expected results and outperformed general deep learning models in terms of performance metrics.

The model fully integrates the structural characteristics of campus wireless networks, leveraging the convenience advantages generated by the partitioned coverage of wireless networks. It collects changes in the number of users directly from the wireless access controllers, thereby generating crowd flow datasets. This approach does not require the deployment of dedicated sensors, nor does it necessitate the active participation of teachers and students. It does not interfere with the daily teaching activities of the school, as all work is completed in the system's background. This provides a new solution for predicting crowd flow in campus environments.

The findings from this research are significant and pave the way for further advancements. Future studies will build on this initial work, aiming to refine and improve the accuracy of crowd flow prediction models. The researchers will draw on wireless positioning technology, a method that has the potential to provide even more precise and detailed data, thereby enhancing the predictive models' capabilities. Moreover, we will collect more campus IoT data, such as campus cards, access control, lighting, and sound, to establish a more comprehensive campus crowd prediction model. By optimizing the model with multi-dimensional feature values, we aim to achieve more accurate crowd flow prediction and precise positioning. In the realm of predictive models, we will endeavor to construct a composite model that is more suitable for predicting the behavior of campus populations. By leveraging the strengths of deep learning, we aim to enhance the accuracy of predictions and improve the efficiency of the algorithms.

References

[1] Z. Chen, L. Zhang, C. Jiang, Z. Cao and W. Cui, "WiFi CSI Based Passive Human Activity Recognition Using Attention Based BLSTM" in IEEE Transactions on Mobile Computing, 18, 11Nov. 2019pp. 2714-2724.

[2] P. Fuxjaeger, S. Ruehrup, T. Paulin and B. Rainer, "Towards Privacy-Preserving Wi-Fi Monitoring for Road Traffic Analysis," in IEEE Intelligent Transportation Systems Magazine, 8, 3, Fall 2016, pp.63-74.

[3] M.Traunmueller, J. Nicholas, M Awais, et al. Digital footprints: Using WiFi probe and locational data to analyze human mobility trajectories in cities," Computers, Environment and Urban Systems, 72, 2018, pp.4-12.

[4] L. W. Chen, Z. G. Chen, C. X. Huang, Y. Q. Lu, J. J. Yang and M. Huang, "Study of the Crowd Behavior in Campus Based on WIFI Probe and Time Series Analysis," 2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science, Rome, Italy, 2020, pp. 1-4.

[5] C. H. Liu et al., "Modeling Citywide Crowd Flows using Attentive Convolutional LSTM," in 37th IEEE International Conference on Data Engineering (IEEE ICDE), Electr Network, Apr 19-22 2021, in IEEE International Conference on Data Engineering, 2021, pp. 217-228.

[6] M. V. Ramesh, S. Anjitha, P. Rekha, and Ieee, "A Novel Wireless Sensor Network Architecture for Crowd Disaster Mitigation," in 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), Shanghai, PEOPLES R CHINA, Sep 21-23 2012.

[7] Yamada, Y. Matsuda, H. Suwa, K. Yastunoto, and S. O. C. Ieee Computer, "Crowd Flow Prediction from Mobile Traces Through Time Series PoI Stay Counts," in International Conference on Smart Computing (SMARTCOMP), Osaka, JAPAN, Jun 29-Jul 02 2024, in International Conference on Smart Computing, 2024, pp. 266-271.

[8] K. Yasufuku and A. Takahashi, "Development of a Real-Time Crowd Flow Prediction and Visualization Platform for Crowd Management," Journal of Disaster Research, vol. 19, no. 2, pp. 248-255, Apr 2024.

[9] Xing, J., Kong, X., Xing, W. et al. STGs: construct spatial and temporal graphs for citywide crowd flow prediction. Appl Intell 52, 12272–12281 (2022).

[10] D. He et al., "A Spatio-Temporal Hybrid Neural Network for Crowd Flow Prediction in Key Urban Areas," Electronics, vol. 12, no. 10, May 2023, Art no. 2255.

[11] H. Go and S. Park, "A study on deep learning model based on global-local structure for crowd flow prediction," Scientific Reports, vol. 14, no. 1, Jun 2024, Art no. 12623.

[12] S. Y. Zhang, B. C. Deng, and D. Q. Yang, "CrowdTelescope: Wi-Fi-positioning-based multi-grained spatiotemporal crowd flow prediction for smart campus," Ccf Transactions on Pervasive Computing and Interaction, vol. 5, no. 1, pp. 31-44, Mar 2023.

[13] C. W. Hu, X. F. Liu, S. Wu, F. Yu, Y. K. Song, and J. Zhang, "Dynamic Graph Convolutional Crowd Flow Prediction Model Based on Residual Network Structure," Applied Sciences-Basel, vol. 13, no. 12, Jun 2023, Art no. 7271.

[14] M. Luca, G. Barlacchi, B. Lepri, and L. Pappalardo, "A Survey on Deep Learning for Human Mobility," Acm Computing Surveys, vol. 55, no. 1, Jan 2023, Art no. 7.

[15] B. F. Yao, S. J. Liu, and L. Wang, "Using Machine Learning Approach to Construct the People Flow Tracking System for Smart Cities," Computer Science and Information Systems, vol. 20, no. 2, pp. 679-700, Apr 2023.

[16] Q. S. Teng, S. Y. Sun, W. D. Song, J. Z. Bei, and C. C. Wang, "Deep Convolutional Neural Network for Indoor Regional Crowd Flow Prediction," Electronics, vol. 13, no. 1, Jan 2024, Art no. 172.