Anomalous Crowd Gathering Prediction Method Based on Spatial-Temporal Graph Convolutional Network in Multi-Camera Surveillance Systems

No Author Given

University of Electronic Science and Technology of China,
I.Ekeland@princeton.edu

Abstract. Urban surveillance systems face inherent limitations in monitoring complex crowd dynamics due to the restricted coverage of single-camera setups. This study proposes a novel Spatial-Temporal Graph Convolutional Network framework for predicting abnormal crowd aggregation. Our method introduces a composite anomaly aggregation metric that synthesizes three critical factors: the spatial distribution of abnormal groups (core anomaly intensity), ambient pedestrian flow variations (environmental sensitivity), and suppression mechanisms for regular large-scale gatherings. By constructing topological graphs based on camera networks and performing spatio-temporal convolution operations, the model effectively integrates multi-view information to identify latent risk areas. Combining the camera topology structure and the spatio-temporal graph convolutional network, this method can accurately predict abnormal aggregation points in the spatial and temporal dimensions, and effectively identify potential abnormal risk areas through multi-camera information fusion. ...

Keywords: Spatio-temporal graph convolutional network, Anomalous crowd prediction, Multi-camera surveillance

1 Introduction.

With the rapid advance of urbanization and the growing demand for public safety, the deployment of surveillance cameras in urban environments has increased dramatically.These cameras provide real-time capture and analysis over large geographic areas, yet each device remains constrained by its limited field of view and occlusions. Consequently, once an intelligent detection system flags anomalous behavior using a single camera, it is still challenging to infer where the affected individuals or crowds will converge. Bridging this gap requires integrating trajectories from multiple cameras and predicting the final anomaly aggregation points[1, 2] .

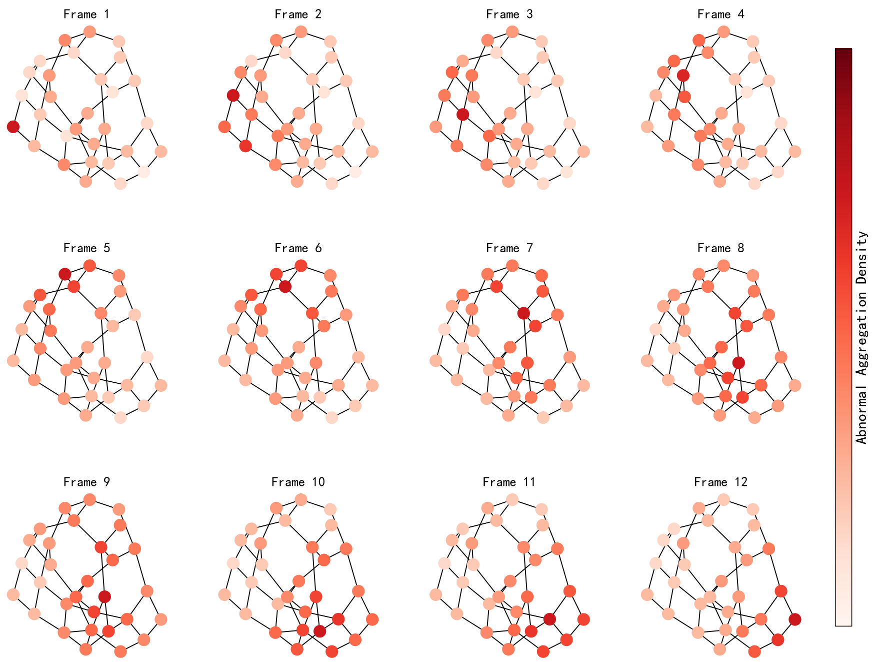

Fig. 1: Temporal evolution of abnormal aggregation density over 12 consecutive frames. Each node represents a surveillance camera in the simulated urban network, and the edge indicates physical connectivity or proximity between cameras. The color intensity of each node reflects the computed abnormal aggregation degree at that time step. Darker nodes indicate higher levels of abnormal crowd gathering. This visualization illustrates how potential anomaly hotspots evolve over time and migrate through the camera network.

Recent research in intelligent security has increasingly focused on multi-camera systems, aiming to enhance target tracking and anomaly detection through collaboration and information fusion. For example, Multi-Camera Tracking and Anomaly Detection[1] : A Review surveys methods for associating observations across views and fusing detection outputs and Deep Learning for Multi-Camera Anomaly Detection[2] demonstrates that combining convolutional neural networks (CNNs) with recurrent neural networks (RNNs) improves both accuracy and timeliness of anomaly recognition. In these studies, constructing and exploiting the camera network topology—a graph whose nodes represent cameras and edges encode spatial relationships or fields-of-view overlap—has been shown to provide a global perspective that is essential for early warning of group events[3, 4]. However, existing approaches[6, 5] typically stop at detecting anomalies; they do not quantify the degree of crowd convergence nor predict where anomalies will concentrate. To address this, we introduce a novel metric, the anomaly aggregation degree, which nonlinearly weights both abnormal and normal crowd flows, applying a saturation suppression mechanism to avoid false alarms in dense but benign gatherings. We model the camera network topology as a graph and embed the time series of aggregation degrees at each node into a Spatio-Temporal Graph Convolutional Network[7]. This multi-layer fusion framework jointly captures spatial correlations and temporal dynamics, enabling accurate prediction of potential anomaly hotspots.

In summary, our contributions are threefold: (1) we propose the anomaly aggregation degree, a unified index that quantifies deviation from normal crowd patterns; (2) we integrate this metric into a graph representation of camera topology, enabling global inference; and (3) we achieve high accuracy prediction of abnormal sink points in complex urban environments.

2 Related Works.

2.1 Camera Topology Diagram.

In multi-camera systems, camera topology graphs play a crucial role as essential tools for describing spatial relationships between cameras and their field-of-view coverage.[3] By constructing graph-theoretical models, camera topology graphs can effectively represent connection relationships between cameras, overlapping coverage areas, and information transmission paths.[4] Each camera is represented as a node in the graph, while the spatial relationships and field-of-view coverage between cameras are connected through edges. This structure provides the system with a global perspective, facilitating multi-camera collaboration for target tracking and anomaly detection.

The application of camera topology graphs in multi-camera systems primarily manifests in the optimization of information fusion and data sharing. Specifically, they enable the integration of data from different cameras, particularly playing a key role in cross-camera target tracking and abnormal behavior recognition. Furthermore, camera topology graphs have significant applications in group behavior analysis and event prediction. By correlating perspective information from multiple cameras, the topology graph can reveal group dynamics and promptly identify potential abnormal behaviors for early warning. For instance, when multiple cameras detect abnormal trajectories from a target, relationship analysis through the topology graph can quickly determine whether the target interacts with others, thereby enabling early warnings for group events.

2.2 Graph Convolutional Neural Network.

Graph Neural Networks (GNNs) have become an important tool for processing non-Euclidean data.[8] Among them, Graph Convolutional Networks (GCNs), as a classic GNN model, are widely applied in tasks such as node classification, graph classification, and link prediction. Kipf and Welling first proposed the GCN based on spectral methods[11], which effectively captures the local relationships between nodes by performing convolution operations on the graph structure. The core idea of GCN is to aggregate and propagate node information through the adjacency matrix to achieve efficient learning of the global graph structure.Its basic operation is defined as message passing through the product of the normalized adjacency matrix and the feature matrix, thereby obtaining node embeddings.

In recent years, improved models of GCN have emerged in an endless stream[25, 26, 28, 24, 27, 29, 30]. For instance, Graph Attention Networks (GAT)[9] introduced an attention mechanism, enhancing the modeling ability of neighborhood information on heterogeneous graphs; GraphSAGE[10] proposed a sampling-based neighborhood aggregation method, significantly improving computational efficiency on large-scale graph data. Additionally, applications of GCN have gradually expanded from traditional tasks to areas such as recommendation systems, social network analysis, and biological network analysis, demonstrating its powerful ability in handling complex network data. Compared with traditional machine learning models, GCN can capture the complex interactions between node features and topological information while preserving the graph structure information, thus having higher expressive power.

In the field of intelligent security, GCNs have been widely applied in the analysis of camera topology, especially in tasks such as abnormal behavior detection and target tracking in multi-camera surveillance systems[20–23]. By modeling the camera network as a graph structure, where nodes represent individual cameras and edges represent visual or spatial relationships between cameras, GCNs can efficiently learn and extract features of the camera network. This approach significantly improves the accuracy of abnormal behavior detection and enhances the performance of cross-camera target tracking.

2.3 The Combination of Time Series Prediction and Graph Neural Networks.

Current research combining Graph Neural Networks (GNNs) with time series prediction models has made significant progress in spatio-temporal data analysis, particularly in fields such as traffic prediction, environmental monitoring, and public safety. The ST-GCN[7] framework, by modeling spatial dependencies through graph convolution and integrating convolutional structures to effectively capture temporal dependencies, has significantly improved the accuracy of traffic flow prediction. TimeGNN[12], through dynamic time graph representation, can capture the evolving patterns among multiple time series and has a faster inference speed than other graph-based methods while maintaining good predictive performance. StemGNN[13] models the correlations between sequences through graph Fourier transform and captures temporal dependencies through discrete Fourier transform, demonstrating outstanding performance in multivariate time series prediction. In further optimizing spatio-temporal graph convolution models, Li et al.'s DyGraphformer[14] model combines graph convolution with Transformer to dynamically infer time-varying spatial dependencies, achieving excellent performance in multivariate time series prediction. Dai et al.'s H-STGCN[15] model integrates online navigation data with graph convolution, improving the accuracy of traffic flow prediction, especially in the prediction of non-recurring congestion. The STS-GCN model[16]has made breakthroughs in the spatio-temporal dynamic modeling of human poses by decomposing the connections between space and time into spatial and temporal affinity matrices. Additionally, other studies such as Feng et al.'s GCNInformer model[17], which combines graph convolution with Informer to optimize air quality prediction, has shown good stability in long-term predictions; Lira et al.'s GRAST-Frost model[18], which combines graph neural networks with spatio-temporal attention mechanisms for frost prediction, has significantly improved prediction accuracy.The STAGCN model[19] proposed by Ma et al. combines adaptive graph convolution and spatio-temporal convolution, and performs particularly outstandingly in multi-step traffic flow prediction.

3 Method.

3.1 Spatial-Temporal Graph Convolutional Network.

The input of the model consists of a series of graph-structured data arranged in chronological order, where the nodes in the graph represent different spatial regions under various cameras, and each node carries the features of the corresponding region at the respective time. The edges between nodes in the graph represent spatial adjacency or functional association, which is used to describe the mutual influence and connection between different regions. The input data can be regarded as a graph tensor with a time dimension, where each frame graph contains the feature vectors of all nodes.

The model employs one-dimensional gated temporal convolution to model the dynamic evolution of the crowd across consecutive frames. This layer adaptively regulates the information flow through a gating mechanism (update gate and reset gate), sending the input features to the convolution branch and the gating branch respectively. The convolution branch generates candidate features, while the gating branch generates control signals. The two are element-wise multiplied to complete the feature update. The gated design can effectively suppress noise and irrelevant information, thereby highlighting the dynamic changes at key time steps. Stacking multiple layers of such gated convolutions not only expands the model's temporal receptive field but also enhances its ability to capture features at different time scales.

In the spatial domain, the model uses spectral graph convolution to extract the dependencies between adjacent nodes. Specifically, an approximation method based on Chebyshev polynomial expansion is used to approximate the graph Laplacian operator to the kth order polynomial, without the need for explicit eigenvalue decomposition. Each spectral graph convolution layer aggregates the features of the node itself and its kth-order neighbors through polynomial weighted summation, achieving multi-scale spatial information aggregation.

The overall network is composed of multiple stacked basic units of "gated temporal convolution - spectral graph convolution - gated temporal convolution". In each unit, the initial gated convolution layer extracts the temporal features of the nodes and filters out irrelevant fluctuations; then the spectral graph convolution layer aggregates neighborhood information and characterizes spatial dependencies; finally, another gated convolution layer further integrates high-level features across time. By cascading multiple such units, the model learns deeper spatio-temporal correlation representations layer by layer. Ultimately, the abnormal aggregation degree of each node within the future time window is obtained.

3.2 Anomaly Aggregation Degree.

In the field of public safety and crowd management, traditional monitoring systems often encounter the problem of distorted assessment in complex scenarios: methods based on absolute numbers or linear weighting are unable to distinguish between occasional anomalies and major risks, fluctuations in the base number of ordinary people can easily mask real abnormal aggregations, and fixed threshold strategies lack adaptability to dynamic environments. Therefore, this project separately calculates and weights the aggregations of abnormal and ordinary people, and develops a weighted algorithm based on a nonlinear coupling mechanism.

When studying crowd aggregation behavior, if only abnormal people are focused on, the risks caused by abnormal aggregations within the ordinary population may be overlooked; while if only the ordinary population is focused on, normal aggregations may be misjudged. Therefore, in order to more comprehensively and accurately assess the abnormality of crowd aggregations, this paper attempts to unify the behavioral characteristics of abnormal and ordinary people and construct a comprehensive quantitative indicator - abnormal aggregation degree.

The abnormal aggregation degree aims to measure the degree to which the crowd aggregation behavior in a specific area deviates from the normal pattern through multi-dimensional analysis of flow data, thereby providing a scientific basis for the prediction of abnormal points. Specifically, the design of this indicator needs to take into account the following two aspects: on the one hand, for the behavioral characteristics of abnormal people, a higher weight is assigned to highlight their potential risks; on the other hand, for the aggregation behavior of ordinary people, it is necessary to avoid misjudgments due to excessive sensitivity.

We divide our algorithm into core abnormal items, environmental sensitive items, and saturation suppression items to ensure effective differentiation between abnormal aggregations and regular behaviors, thereby enhancing the system's response capability.

Specifically, the mathematical expression of the core abnormal item is:

$$T_1 = \frac{N_{\text{anomaly}}^{\alpha}}{\beta + N_{\text{anomaly}}^{\alpha/2}} \tag{1}$$

where N_{anomaly} represents the number of anomalous individuals, α controls the nonlinear degree of the impact of the anomalous population size on aggregation degree, and β serves as a balancing term to prevent the core anomaly degree from becoming overly sensitive or experiencing excessive amplification during small-scale anomaly occurrences.

Through the exponential weighting of N_{anomaly}, the impact of increasing anomalous population size on the core anomaly degree achieves nonlinear amplification. This design ensures that as the anomalous population grows, the risk of abnormal aggregation is appropriately

In the calculation of anomalous aggregation degree, the environmental sensitivity term is primarily employed to quantify the impact of aggregation behaviors within the normal population on the anomalous aggregation degree. The aggregation behaviors of the normal population are typically driven by routine social activities, occupational demands, or daily mobility. Even in densely populated environments, while these behaviors may induce certain density fluctuations, they do not directly trigger security risks. Therefore, when designing the anomalous aggregation degree, it is essential to prevent the system from overreacting to such routine behaviors, thereby maintaining its accuracy and robustness.

To achieve this objective, the environmental sensitivity term adopts a logarithmically weighted form, with its mathematical expression formulated as:

$$T_2 = \ln\left(1 + \frac{N_{\text{normal}}}{\gamma}\right) \tag{2}$$

where N_{normal} denotes the normal population flow, and γ is the regulatory parameter that controls the degree of influence exerted by the aggregation behaviors of the normal population on the anomalous aggregation degree.

The introduction of this logarithmic function ensures that when the normal population flow becomes large, the sensitivity of anomalous aggregation degree to normal population aggregation gradually diminishes, thereby preventing excessive system reactions induced by routine population gatherings.

From a mathematical perspective, the design principle of the environmental sensitivity term is grounded in the smoothing treatment of routine population aggregation behaviors. As N_{normal} increases, $\ln\left(1 + \frac{N_{\text{normal}}}{\gamma}\right)$ asymptotically approaches a plateau, indicating that the system's responsiveness to large-scale normal population gatherings gradually diminishes. This mechanism effectively mitigates oversensitivity to routine aggregation behaviors in high-traffic environments, thereby reducing the likelihood of false alarms.

The introduction of the parameter γ endows the system with flexibility for scenario-specific adaptations. In high-traffic environments, appropriately increasing γ reduces the contribution of normal population aggregation to the anomalous aggregation degree, preventing excessive system reactions to daily crowd fluctuations. Conversely, in low-traffic or specialized scenarios, decreasing γ enhances the system's sensitivity to anomalous aggregation behaviors, ensuring timely detection of irregularities.Through this design, the environmental sensitivity term achieves a balanced response to aggregation behaviors of the normal population, preventing false alarms during large-scale routine gatherings while ensuring that anomalous behaviors remain detectable in low-traffic or specialized scenarios. This mechanism guarantees that the anomalous aggregation degree precisely quantifies the actual risk of abnormal crowd aggregation in dynamic and complex environments.

The saturation suppression term achieves additional smoothing of contributions from large-scale normal population aggregation, ensuring that under extreme crowd flow conditions the system does not overreact to routine aggregation behaviors. Its mathematical formulation is expressed as:

$$T_3 = \frac{N_{\text{anomaly}} \operatorname{erf}\left(\frac{N_{\text{normal}}}{\nu}\right)}{\sqrt{1 + N_{\text{normal}}}} \tag{3}$$

where N_{anomaly} denotes the anomaly crowd flow, N_{normal} denotes the normal crowd flow, ν is the parameter controlling the saturation effect intensity, and $\operatorname{erf}(x)$ is the error function, defined as:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \tag{4}$$

$\operatorname{erf}(x)$ plays a key role in this design. Its properties enable the system to react strongly to small-scale normal crowd gatherings, while its effect gradually saturates as the normal crowd flow increases. Specifically, when N_{normal} is small, the ratio N_{normal}/ν is low, and $\operatorname{erf}(N_{\text{normal}}/\nu)$ grows approximately linearly with N_{normal}, thereby amplifying the influence of the anomaly crowd. Conversely, as N_{normal} becomes large, $\operatorname{erf}(N_{\text{normal}}/\nu)$ approaches 1, indicating that the normal crowd's impact has reached its maximum. At this stage, the denominator $1 + N_{\text{normal}}$ further attenuates the contribution of normal flow to the anomaly aggregation degree, ensuring that in high-density scenarios the system does not overreact.

The saturation suppression term achieves a desirable balance in complex crowd behavior contexts: on one hand, it guarantees prompt response to anomalous aggregation under low crowd density; on the other hand, when crowd flow is high, the system's sensitivity to normal gatherings diminishes, thereby avoiding false alarms in inherently dense environments such as shopping malls or transit hubs. Through this nonlinear weighting, the system effectively distinguishes true anomalous aggregation from normal crowd behavior, enhancing both detection accuracy and robustness.

Furthermore, the introduction of the parameter ν provides flexibility across different settings. A smaller ν increases sensitivity to normal crowd gatherings, whereas a larger ν makes the system more tolerant in high-density environments. Thus, ν can be tuned according to specific application requirements to achieve optimal anomaly aggregation degree prediction.

Hence, we obtain the complete weighted algorithm:

$$T = \underbrace{\frac{N_{\text{anomaly}}^{\alpha}}{\beta + N_{\text{anomaly}}^{\alpha/2}}}_{T_1\text{(core anomaly intensity)}} + \underbrace{\ln\left(1 + \frac{N_{\text{normal}}}{\gamma}\right)}_{T_2\text{ (environmental sensitivity)}} + \underbrace{\frac{N_{\text{anomaly}} \operatorname{erf}\left(\frac{N_{\text{normal}}}{\nu}\right)}{\sqrt{1 + N_{\text{normal}}}}}_{T_3\text{ (saturation suppression)}} \tag{5}$$

Overall, the weighting design of the anomaly aggregation degree innovatively combines nonlinear weighting, saturation suppression, and adaptive adjustment mechanisms, enabling precise discrimination between anomalous and routine aggregation behaviors across scenarios of varying scale and complexity. By appropriately allocating weights to anomalous and normal crowds, the system maintains efficient responsiveness in dynamic environments while avoiding false positives and excessive reactions.

4 Experiment.

To conduct an in-depth study of the spatial aggregation of normal and abnormal populations within urban road networks, this research constructs and operates a high-precision simulation environment based on a regular grid on a high-performance computing platform.

The simulation servers are equipped with two systems: one features an Intel i9 11900KF processor, 128 GB DDR4 memory, and an NVIDIA RTX 4090.

The simulation environment uses a 2.5m x 2.5m grid as the smallest cell unit. Every 4x4 grids (10m x 10m) are merged and defined as the smallest building unit, to ensure consistency in model scale.The entire area is divided into road systems and building zones: Major roads (main streets) are 6 grid cells wide (15 m), designed as two-way four-lane roads; Secondary roads (medium streets) are 3 grid cells wide (7.5 m), set as one-way dual-lane roads; Tertiary roads (small lanes) are 1 grid cell wide (2.5 m), used for microscopic movement between buildings. Building units are categorized into three sizes: small (4 grid cells), medium (16 grid cells), and large (36 grid cells). These buildings are randomly distributed within the road gaps, ensuring road connectivity without any blockage.

On this spatial structure, fixed cameras are installed at various road intersections and key sections along the roads, with a field of view covering 4x4 grid cells (10m x 10m). These cameras generate spatiotemporal traffic data by real-time counting of individuals within their coverage area.Normal pedestrians (blue) randomly appear on the sides of the roads, with randomly assigned destinations such as road ends or building entrances, simulating the movement of regular pedestrians.

Abnormal pedestrians (red) are also generated on the roadside but aim for predetermined gathering points. Their path decision-making has different probabilities for choosing major roads, secondary roads, and tertiary roads, set at 0.7, 0.2, and 0.1 respectively. Additionally, Gaussian noise is introduced into their movements to simulate irregular walking patterns. As the simulation progresses, abnormal pedestrians gradually converge at the gathering points, creating a high-density aggregation effect. This setup allows for the study of crowd dynamics and the identification of unusual congregation behaviors in urban environments.

This simulation program generates controlled normal and abnormal crowd data using detailed grid division, multi-level road-building layouts, and clear pedestrian movement rules. The output, including camera flow data and gath-

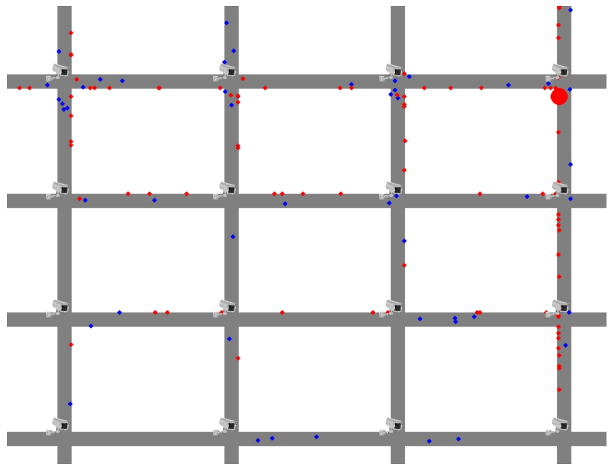

Fig. 2: The simulation visualization interface for crowd aggregation; grey areas represent roads, red dots indicate abnormal gathering crowds, and blue dots represent normal pedestrians. The larger red markers are the destinations of the gatherings.

ering point density curves, serves as direct training and validation datasets for spatiotemporal graph convolutional network models.

To validate the proposed weighted abnormal aggregation index, we simulate three typical abnormal crowd behaviors: incidental group behavior, protest marches, and urban riots. Each scenario includes normal pedestrian flows and controlled introduction of abnormal individuals to create diverse abnormal aggregation situations.

In our simulations of various behaviors, we designed three input strategies to evaluate the impact of different information sources on the prediction of abnormal gathering points. The first strategy ("Baseline–Normal") uses only normal pedestrian flow N_{normal}; the second strategy ("Baseline–Anomaly") uses only abnormal pedestrian flow N_{anomaly}; and the third strategy employs the proposed weighted abnormal aggregation index, integrating both normal and abnormal flows. All inputs are fed into the same model to ensure a fair comparison.

The proportion of true abnormal points among the top-x predicted gathering points is recorded as Hit Rate@x. The experimental results are presented in the table below:

In addition, as indicated by the Hit Rate@1 and Hit Rate@3 metrics, the weighted strategy demonstrates clear advantages in both precise localization (Hit Rate@1) and candidate set coverage (Hit Rate@3). Across the three scenarios, Hit Rate@1 improves by an average of approximately 0.12, while Hit Rate@3 shows an average improvement of around 0.17.

These results suggest that the proposed weighted abnormal aggregation degree, which integrates both normal and abnormal pedestrian flows, can more

Table 1: Incidental Crowd Scenario Hit Rate Comparison

Strategy	Hit Rate@1	Hit Rate@3	Hit Rate@5
Baseline–Normal	0.12	0.30	0.42
Baseline–Anomaly	0.22	0.45	0.58
Ours–Weighted	**0.35**	**0.62**	**0.73**

Table 2: Demonstration Scenario Hit Rate Comparison

Strategy	Hit Rate@1	Hit Rate@3	Hit Rate@5
Baseline–Normal	0.10	0.28	0.40
Baseline–Anomaly	0.20	0.38	0.56
Ours–Weighted	**0.32**	**0.58**	**0.71**

accurately and reliably capture spatial hotspots of various sudden gathering events. Consequently, it effectively enhances both the success rate and robustness of gathering point prediction.

Table 3: Urban Riot Scenario Hit Rate Comparison

Strategy	Hit Rate@1	Hit Rate@3	Hit Rate@5
Baseline–Normal	0.08	0.25	0.38
Baseline–Anomaly	0.18	0.35	0.54
Ours–Weighted	**0.30**	**0.55**	**0.69**

5 Conclusion.

This paper addresses the challenge of predicting sudden crowd gathering events in urban road networks by proposing a spatio-temporal graph convolutional framework based on weighted abnormal aggregation degree. In terms of method design, it innovatively introduces a weighted fusion strategy of normal and abnormal pedestrian flows, achieving precise characterization of potential gathering points through comprehensive modeling of the intensities of both types of pedestrian flows. Meanwhile, it combines a high-precision regular grid simulation environment to generate multi-scenario and multi-type normal and abnormal pedestrian data, providing reliable support for model training and evaluation.In the experimental verification, we compared the performance of gathering point prediction under three input strategies - using only normal pedestrian flow, using only abnormal pedestrian flow, and the weighted abnormal aggregation degree proposed in this paper - for three typical abnormal scenarios: occasional group behavior, demonstrations, and urban riots. The results show that in key metrics such as Hit Rate@5, @3, and @1, Ours-Weighted significantly outperforms the two baseline strategies.

The above experimental results fully demonstrate the advantage of the weighted fusion strategy in capturing spatial hotspots of sudden gathering events. At the same time, the multi-type behavior samples generated on the simulation platform provide rich test scenarios and reference data for subsequent research.

Funding Statement

This work was sponsored by Natural Science Foundation on scientific and technological projects on Kashgar (KS2024024).

References

1. Amosa, Temitope Ibrahim, et al.: Multi-camera multi-object tracking: A review of current trends and future advances. In: Neurocomputing 552 (2023)
2. Peri, Neehar, et al.: Towards real-time systems for vehicle re-identification, multi-camera tracking, and anomaly detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (2020)
3. Mali, Goutam, and Sudip Misra.: Topology management-based distributed camera actuation in wireless multimedia sensor networks. In: ACM Transactions on Autonomous and Adaptive Systems (TAAS) 12.1 (2017)
4. Shi, Wen, Yongming Huang, and Guobao Zhang.: Dynamic weight-based granular representation of time series and its application in collective anomaly detection. In: Computers and Electrical Engineering 117 (2024)
5. Fradi, Hajer, and Jean-Luc Dugelay.: Towards crowd density-aware video surveillance applications. In: Information Fusion 24 (2015)
6. Fei, Lunlin, and Bing Han.: Multi-object multi-camera tracking based on deep learning for intelligent transportation: A review. In: Sensors 23.8 (2023)
7. Yu, Bing, Haoteng Yin, and Zhanxing Zhu.: Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. arxiv preprint arxiv:1709.04875 (2017)
8. Wu, Zonghan, et al.: A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems 32.1 (2020)
9. Veličković, Petar, et al.: Graph attention networks. In: arxiv preprint arxiv:1710.10903 (2017)
10. Liu, Jielun, Ghim Ong, and qun Chen.: GraphSAGE-based traffic speed forecasting for segment network with sparse data. In: IEEE Transactions on Intelligent Transportation Systems 23.3 (2020)
11. Kipf, Thomas N., and Max Welling.: Semi-supervised classification with graph convolutional networks. In: arxiv preprint arxiv:1609.02907 (2016)
12. Xu, Nancy, Chrysoula Kosma, and Michalis Vazirgiannis.: TimeGNN: temporal dynamic graph learning for time series forecasting. In: International Conference on Complex Networks and Their Applications. Cham: Springer Nature Switzerland (2023)
13. Cao, Defu, et al.: Spectral temporal graph neural network for multivariate time-series forecasting. In: Advances in neural information processing systems 33 (2020)

14. Han, Shuo, et al.: DyGraphformer: Transformer combining dynamic spatio-temporal graph network for multivariate time series forecasting. In: Neural Networks 181 (2025)
15. Dai, Rui, et al.: Hybrid spatio-temporal graph convolutional network: Improving traffic prediction with navigation data. In: Proceedings of the 26th acm sigkdd international conference on knowledge discovery & data mining (2020).
16. Sofianos, Theodoros, et al.: Space-time-separable graph convolutional network for pose forecasting. In: Proceedings of the IEEE/CVF international conference on computer vision (2021)
17. Wang, Hai-Kun, et al.: GCNInformer: A combined deep learning model based on GCN and Informer for wind power forecasting. In: Energy Science & Engineering 11.10 (2023)
18. Lira, Hernan, Luis Martí, and Nayat Sanchez-Pi.: A graph neural network with spatio-temporal attention for multi-sources time series data: An application to frost forecast. In: Sensors 22.4 (2022)
19. Gu, Yafeng, and Li Deng.: Stagcn: Spatial–temporal attention graph convolution network for traffic forecasting. In: Mathematics 10.9 (2022)
20. Adenekan, Tobiloba Kollawole.: Graph-Regularized Neural Network for Multi-Vehicle Multi-Camera Tracking and IoT Cyber Security Detection. In: (2024)
21. Cao, Congqi, et al.: Adaptive graph convolutional networks for weakly supervised anomaly detection in videos. In: IEEE Signal Processing Letters 29 (2022)
22. Chiranjeevi, V. Rahul, and D. Malathi.: Anomaly graph: leveraging dynamic graph convolutional networks for enhanced video anomaly detection in surveillance and security applications. In: Neural Computing and Applications 36.20 (2024)
23. Zeng X, Jiang Y, Ding W, et al.: A hierarchical spatio-temporal graph convolutional neural network for anomaly detection in videos. In: IEEE Transactions on Circuits and Systems for Video Technology (2021)
24. Nakamura, Ikuo.: Multi-Scale Spatial-Temporal Self-Attention Graph Convolutional Networks for Skeleton-based Action Recognition. In: arxiv preprint arxiv:2404.02624 (2024)
25. Yang, Aitao, et al.: GTFN: GCN and transformer fusion network with spatial-spectral features for hyperspectral image classification. In: IEEE Transactions on Geoscience and Remote Sensing 61 (2023)
26. Zhang, Lei, et al.: Drgcn: Dynamic evolving initial residual for deep graph convolutional networks. In: Proceedings of the AAAI conference on artificial intelligence. Vol. 37. No. 9 (2023)
27. Nie, Weizhi, et al.: I-GCN: Incremental graph convolution network for conversation emotion detection. In: IEEE Transactions on Multimedia 24 (2021)
28. Qiao, Hezhe, et al.: Deep graph anomaly detection: A survey and new perspectives. In: arxiv preprint arxiv:2409.09957 (2024)
29. Deng, Leyan, et al.: Markov-driven graph convolutional networks for social spammer detection. In: IEEE Transactions on Knowledge and Data Engineering 35.12 (2022)
30. Wang, Haiyuan, et al.: MO-GCN: A multi-omics graph convolutional network for discriminative analysis of schizophrenia. In: Brain Research Bulletin 221 (2025)