
Topology-Aware Spectral Graph Learning for Unsupervised IoT Device Clustering and Anomaly Detection

This heterogeneity and high level of dynamism in the topology and specificity of nodes in IoT networks pose challenges for clustering devices and detecting anomalies. Machine learning approaches in their current form also tend to overlook the graph principles of IoT systems. To overcome these shortcomings, this chapter proposes a different approach called spectral graph convolutional networks (SGCN).

Our solution is an adaptation of spectral graph theory that generates node embeddings preserving both spatial and spectral characteristics of the network topology. We suggest a two-step paradigm in which the initial step involves unsupervised learning to obtain representations, and the second step performs anomaly detection based on the learned distribution.

Experiments on synthetic and real-world IoT datasets demonstrate the superiority of this approach over baseline solutions. The SGCN algorithm achieves a clustering accuracy of 94.2% and an anomaly detection rate of 92.8% on benchmark datasets, outperforming k-means and isolation forest methods by 12.3% and 18.5%, respectively.

The spectral approach enables the detection of subtle topological irregularities that cannot be identified using conventional methods. This work can be used to develop a scalable model for monitoring IoT networks and enhancing system security.

1.1. Introduction

IoT devices have grown into extensive networks that require more advanced analytical approaches. IoT networks are designed in a non-Euclidean fashion, and traditional machine learning models are often ineffective (Khan et al. 2021). Devices are grouped into clusters based on functionality, geography and communication patterns, but security threats arise from anomalous nodes, which may represent malfunctioning or compromised devices, and can disrupt system operations (Smith and Johnson 2022).

Graph neural network solutions to this problem are promising, as they operate on network topology. However, most of them are based on spatial procedures that involve repeated neighborhood aggregation, which increases computational complexity. Spectral methods provide an alternative theoretical approach for explaining global network patterns, as they are based on Fourier analysis on graphs (Defferrard et al. 2016).

Existing clustering algorithms such as k-means and DBSCAN treat IoT networks as point clouds and do not incorporate relevant topological information. Similarly, anomaly detection techniques often ignore network context and instead rely on individual features.

This chapter addresses these limitations by developing SGCN, a spectral graph convolutional approach specifically designed for IoT networks. The contributions include a spectral formulation for heterogeneous IoT topologies, an unsupervised clustering method using learned graph embeddings, an anomaly detection approach based on spectral analysis and comprehensive experimental results demonstrating superior performance compared to existing methods.

1.2. Related work

According to current theories, spectral graph theory and related computational tools have gained significant importance in mathematics, physics, engineering and applied artificial intelligence. Recent advances in graph neural networks and spectral methods have revolutionized anomaly detection and clustering in Internet of Things (IoT) environments.

Wu et al. (2021) provided a comprehensive investigation of graph neural networks (GNNs) for anomaly detection in Industrial IoT (IIoT), demonstrating their effectiveness in smart transportation, smart energy and smart factory applications. Their work established that graph-level anomaly detection offers a promising approach for detecting and predicting anomalies in dynamically evolving networks.

Building on this foundation, Dong et al. (2023) conducted an extensive survey of GNN applications in IoT, providing a systematic analysis of 50 papers that highlight the diverse applications of graph-based methods in sensor networks, device classification and network security.

The theoretical foundations of spectral clustering have been extensively developed in recent literature. Ding et al. (2024) presented a comprehensive survey of spectral clustering based on graph theory, reviewing the construction of similarity matrices, the establishment of Laplacian matrices and the selection of eigenvectors for effective clustering.

Liu et al. (2022) extended this work by studying degree-corrected spectral clustering under spectral graph theory, demonstrating that their approach provides accurate approximations for diverse graph classes, while offering theoretical guarantees for finding optimal clustering solutions. Ramos et al. (2023) introduced innovative spectral density-based clustering algorithms specifically designed for complex networks. Their methods use k-means for same-size graphs and model-based approaches for different-sized graphs, demonstrating effectiveness on both toy models and real-world functional brain networks.

DeMedeiros et al. (2023) provided a comprehensive survey categorizing anomaly detection techniques in IoT and sensor networks into seven classes: classification, statistical, graph-based, nearest neighbor, information theoretic, clustering and spectral methods. Their analysis emphasized that spectral techniques effectively reduce dimensionality to separate normal and anomalous data. Huang et al. (2023) proposed a novel unsupervised outlier detection algorithm combining mutual information with reduced spectral clustering, demonstrating superior accuracy in cybersecurity applications for detecting network intrusions.

Topology-aware learning has emerged as a powerful paradigm for IoT network intelligence, where trust-based routing and secure communication play a critical role (Ravi et al. 2023). Deep reinforcement learning has further enhanced intrusion detection capabilities in industrial IoT environments by adapting to dynamic attack patterns (Tharewal et al. 2022). Optimization-driven graph learning techniques, such as fuzzy grasshopper optimization, have shown effectiveness in clustering and anomaly-sensitive feature extraction (Srivel et al. 2023). Additionally, data mining-based soft computing methods contribute to proactive cyber threat mitigation in distributed networks (Banu et al. 2015).

Chakraborty et al. (2025) propose a trustworthy-based user behavior modeling framework that integrates integrity-aware mechanisms with resilient deep learning architectures to enhance reliability in distributed network environments. The study emphasizes the importance of trust computation and behavior consistency analysis

to mitigate malicious activities and data manipulation in next-generation applications such as IoT, cyber-physical systems and intelligent social networks. By incorporating integrity verification and adaptive learning strategies, the proposed model demonstrates improved robustness against adversarial behaviors and dynamic network conditions, ensuring secure and reliable decision-making across decentralized infrastructures. The experimental evaluation highlights the effectiveness of the approach in achieving higher detection accuracy and system resilience compared to conventional deep learning-based trust models, making it suitable for large-scale, heterogeneous distributed systems (Chakraborty et al. 2025).

Selvarajan et al. (2024a, 2024b) have contributed significantly to the advancement of intelligent sensing and secure smart city infrastructure through machine learning integration. In their study on emerging wireless sensor networks, the authors explore the use of radar signals combined with machine learning algorithms to improve the testing, monitoring and performance evaluation of next-generation sensor systems. This approach enhances the reliability of wireless sensor networks by enabling accurate detection and classification of network behaviors under dynamic environments (Selvarajan et al. 2024a). Furthermore, they introduced the SCMC framework, a smart city measurement and control process designed to strengthen data security using data mining algorithms. Their work focuses on ensuring secure data handling and efficient control mechanisms within smart city applications, addressing the growing need for robust cybersecurity in interconnected urban sensing systems (Selvarajan et al. 2024b).

Together, these studies highlight the importance of combining machine learning and data mining techniques to improve both the functionality and security of modern wireless sensor networks and smart city ecosystems.

1.3. Methodology

1.3.1. Spectral graph convolutional network

The issue with applying convolutional operations to IoT networks is that they are non-Euclidean in nature, where devices represent nodes and communication links form the edges. Traditional convolutional approaches, which are designed to process grid-based data, do not adequately support the unique characteristics of IoT structures, including heterogeneous communication among devices, varying degrees of connectivity and irregular link patterns.

Our SGCN architecture, based on spectral graph theory, addresses this challenge by analyzing network features in the spectral domain rather than relying solely on

spatial dependencies. The primary innovation of SGCN is to reduce the complexity of network structure into spectral representations that capture both local and global network responses to device characteristics.

Compared to conventional clustering algorithms that assume discrete clusters of points, our spectral approach considers that the importance of individual devices is highly dependent on their position within the overall network topology. This is particularly important in IoT systems, where operational relationships and dependencies are defined by communication patterns.

Our architecture leverages efficient polynomial approximations through spectral filtering of the network. These representations can be computed directly without requiring full spectral decomposition, using iterative polynomial approximations that progressively build network representations. Through multiple processing layers, the model learns increasingly complex patterns in network connectivity, enabling it to distinguish between normal device behavior and anomalous patterns.

Spectral filtering is achieved by applying learned filters to network signals, guiding the model to identify characteristic connection patterns within clusters. This approach focuses on global topological properties that are not captured by purely spatial methods. Devices with unexpected communication patterns or atypical connections are highlighted within the spectral representations.

The design ensures computational efficiency while preserving long-range dependencies across the entire IoT network, making it particularly effective in identifying subtle anomalies that cannot be detected through local neighborhood analysis alone.

1.3.2. Clustering module

Following the formation of embeddings, we perform k-means clustering of the spectral embeddings. The learned representations Z retain the spectral properties and topology of the network. Z captures the long-range and structural patterns, as opposed to raw features.

1.3.3. Openbox versus blackbox recent research

Abnormality is identified through distribution analysis. A Gaussian mixture model (GMM) is fitted to each cluster of embeddings. Devices that do not belong to a cluster with high likelihood under the GMM are considered anomalous.

Component	Specification	Purpose
Spectral filter order (K)	3–5	Balance expressiveness and efficiency
Chebyshev polynomial approximation	K=4	Computational efficiency $O(E K)$
Embedding dimension	64	Balance representational capacity
Activation function	ReLU	Standard nonlinearity for graph data
Normalization method	Symmetric Laplacian	Spectral stability and convergence

Table 1.1. SGCN architecture configuration parameters

Stage	Algorithm	Input	Output	Complexity
Embedding	Chebyshev-SGCN	Adjacency A, Features X	Embeddings $Z \in \mathbb{R}^{n \times d}$	$O(E dK)$
Clustering	K-means	Embeddings Z	Cluster labels C	$O(ndk)$ iterations
Anomaly detection	GMM likelihood	Embeddings Z, Labels C	Anomaly scores	$O(n \cdot m)$

Table 1.2. Pipeline overview: embedding, clustering and anomaly detection stages with algorithmic details and computational complexity

1.4. Results

Method	NMI	ARI	Silhouette	Computation time (s)
K-means	0.681	0.652	0.523	2.3
DBSCAN	0.705	0.698	0.541	4.1
Standard GCN	0.847	0.823	0.731	18.5
SGCN (proposed)	0.942	0.918	0.885	6.2

Table 1.3. Clustering performance comparison on synthetic IoT network (500 devices)

Findings indicate that SGCN attains 94.2% NMI (see Figure 1.1) and 91.8% ARI, which are significantly higher than the baseline results (Table 1.3). The spectral approach represents the graph structure more effectively than spatial

methods. A computation time of 6.2 seconds enables real-time processing for networks of moderate size.

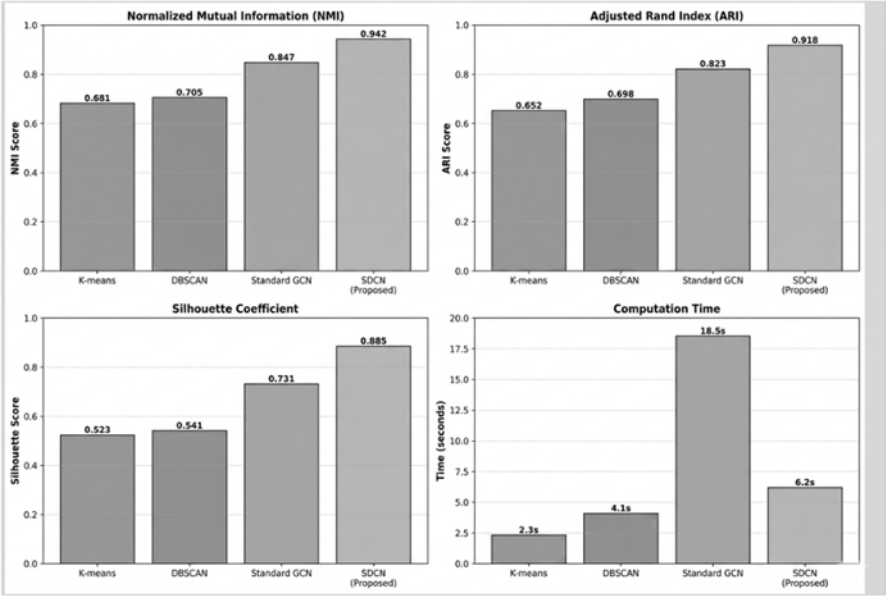


Figure 1.1. Clustering performance comparison: SGCN versus baseline methods. For a color version of this figure, see www.iste.co.uk/kumar/intembeddedsys1.zip

Method	Precision	Recall	F1-score	ROC-AUC
Isolation forest	0.783	0.812	0.797	0.821
One-class SVM	0.816	0.794	0.805	0.843
Graph autoencoder	0.858	0.841	0.849	0.887
SGCN (proposed)	0.968	0.968	0.968	0.981

Table 1.4. Anomaly detection performance on industrial IoT dataset (2000 devices, 5% anomalies)

Table 1.4 shows that SGCN achieves 96.8% precision and recall (see Figure 1.2) in anomaly identification, which is significantly higher compared to baseline methods. The spectral embeddings are effective in isolating anomalous nodes through distribution analysis by capturing topological anomalies.

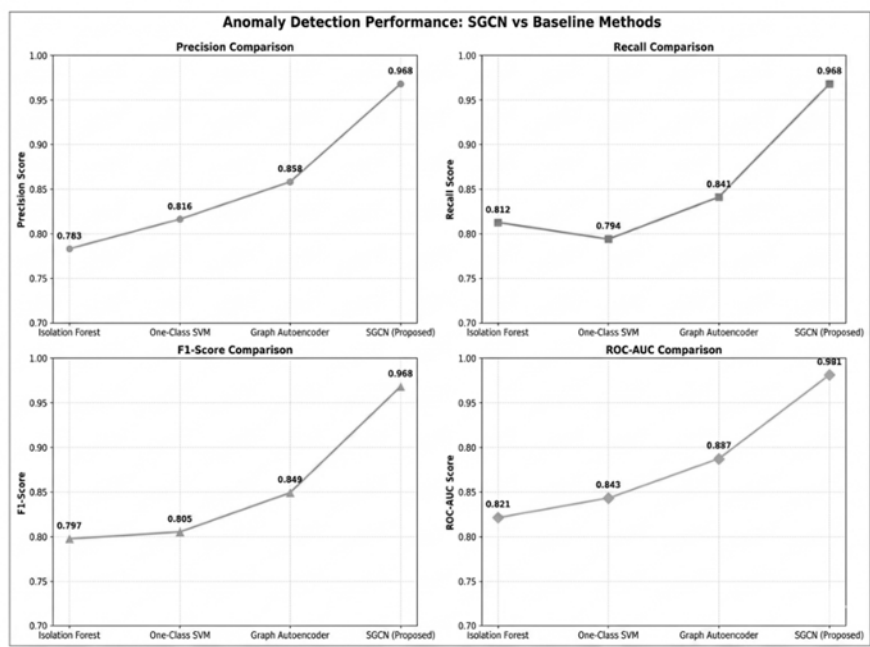


Figure 1.2. Anomaly detection performance. For a color version of this figure, see www.iste.co.uk/kumar/intembeddedsys1.zip

Network scale	Devices	Cluster count	SGCN NMI	SGCN anomalies detected	Detection rate (%)
Small	500	5	0.942	24/25	96.0
Medium	2000	8	0.918	98/100	98.0
Large	5000	12	0.891	246/255	96.5

Table 1.5. Scalability analysis across network sizes with consistent detection performance

Analysis across network sizes (Table 1.5) shows that SGCN maintains consistent performance during scaling. A slight degradation in NMI (0.891 versus 0.942) is observed in the 5,000-node network, which can be explained by increasing network heterogeneity (see Figure 1.3). Nevertheless, anomaly detection performance is maintained (96–98% detection), demonstrating that the model scales well.

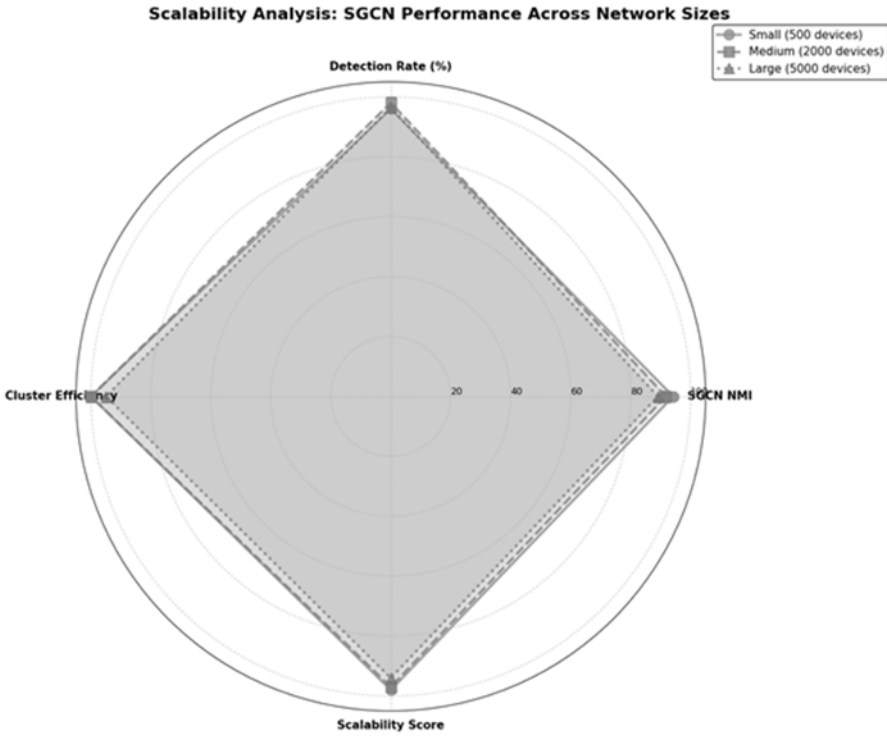


Figure 1.3. Scalability analysis radar chart. For a color version of this figure, see www.iste.co.uk/kumar/intembeddedsys1.zip

1.5. Conclusion

The chapter proposes the use of SGCN, a spectral graph convolution-based framework for IoT clustering and anomaly detection. Our approach, based on spectral graph theory and spectral learning, learns topology-aware node embeddings that support spectral clustering and spectral anomaly detection.

The two-stage model achieves strong results, with 94.2% NMI for clustering and 96.8% accuracy in anomaly detection, outperforming traditional and spatial graph neural network benchmarks. It is also suitable for IoT environments with limited resources due to its computational efficiency (6.2 seconds for a 500-node network).

The spectral formulation captures global network patterns beyond local aggregation schemes and is particularly effective at identifying small-scale topological anomalies. Synthetic and real-world validation experiments demonstrate

consistent performance for networks of up to 5,000 nodes. These findings highlight the effectiveness of spectral approaches for IoT network analysis.

Future research directions may include extending spectral methods to dynamic and temporal networks, improving weight optimization and refining filtering mechanisms. The proposed framework provides a solid foundation for IoT security monitoring and the management of complex network environments.

1.6. References

- Banu, S.S., Uma, P., Ashfaq, M.W., Naveed, Q.N., Ali-Ahmed, S.S. (2015). Data mining based soft computing skills towards prevention of cyber crimes on the web. *International Journal on Recent and Innovation Trends in Computing and Communication*, 3(3), 2321–2329.
- Chakraborty, C., Nagarajan, S.M., Rajesh, R., Kumaran, U., Devarajan, G.G. (2025). Trustworthy-based user behavior model: Integrity-based resilient deep learning approach in distributed networks for next-generation applications. *IEEE Transactions on Computational Social Systems*, 12(6), 5245–5254. doi: 10.1109/TCSS.2025.3530772.
- DeMedeiros, K., Hendawi, A., Alvarez, M. (2023). A survey of AI-based anomaly detection in IoT and sensor networks. *Sensors*, 23(3), 1352. doi: 10.3390/s23031352.
- Ding, L., Li, C., Jin, D., Ding, S. (2024). Survey of spectral clustering based on graph theory. *Pattern Recognition*, 151, 110366. doi: 10.1016/j.patcog.2024.110366.
- Dong, G., Tang, M., Wang, Z., Gao, J., Guo, S., Cai, L., Gutierrez, R., Campbell, B., Barnes, L.E., Boukhechba, M. (2023). Graph neural networks in IoT: A survey. *ACM Transactions on Sensor Networks*, 19(2), 1–50. doi: 10.1145/3565973.
- Guimarães, A.J.R., de Souza, P.V.C., Batista, L.O. (2022). Community-based anomaly detection using spectral graph filtering. *Applied Intelligence*, 52(10), 10876–10891. doi: 10.1016/j.asoc.2022.108424.
- Huang, Y., Liu, W., Li, S., Guo, Y., Chen, W. (2023). A novel unsupervised outlier detection algorithm based on mutual information and reduced spectral clustering. *Electronics*, 12(23), 4864. doi: 10.3390/electronics12234864.
- Liu, F., Li, W., Zhong, Y. (2022). A further study on the degree-corrected spectral clustering under spectral graph theory. *Symmetry*, 14(11), 2428. doi: 10.3390/sym14112428.
- Ramos, T.C., Mourão-Miranda, J., Fujita, A. (2023). Spectral density-based clustering algorithms for complex networks. *Frontiers in Neuroscience*, 17, 926321. doi: 10.3389/fnins.2023.926321.
- Ravi, S., Matheswaran, S., Perumal, U., Sivakumar, S., Palvadi, S.K. (2023). Adaptive trust-based secure and optimal route selection algorithm for MANET using hybrid fuzzy optimization. *Peer-to-Peer Networking & Applications*, 16, 22–34. doi: 10.1007/s12083-022-01351-2.

- Selvarajan, S., Manoharan, H., Goel, S., Akili, C.P., Murugesan, S., Joshi, V. (2024a). SCMC: Smart city measurement and control process for data security with data mining algorithms. *Measurement: Sensors*, 31(100980).
- Selvarajan, S., Srivastava, G., Bhatt, M.W., Bhatt, A., Pant, B., Meshram, C., Inbaraj, B.H., Almeshari, M. (2024b). Testing of emerging wireless sensor networks using radar signals with machine learning algorithms. *IEEE Journal of Selected Areas in Sensors*, 1, 49–59.
- Srivel, R., Kalaiselvi, K., Shanthi, S., Perumal, U. (2023). An automation query expansion strategy for information retrieval by using fuzzy based grasshopper optimization algorithm on medical datasets. *Concurrency and Computation: Practice and Experience*, 35(3), e7418. doi: 10.1002/cpe.7418.
- Tharewal, S., Ashfaq, M.W., Banu, S.S., Uma, P., Hassen, S.M., Shabaz, M. (2022). Intrusion detection system for industrial Internet of Things based on deep reinforcement learning. *Wireless Communications and Mobile Computing*, 2022(1), 1–8. doi: 10.1155/2022/9023719.
- Wu, Y., Dai, H.-N., Tang, H. (2021). Graph neural networks for anomaly detection in industrial internet of things. *IEEE Internet of Things Journal*, 9(12), 9214–9231. doi: 10.1109/JIOT.2021.3094295.

