COLOMBO SCIENTIFIC PUBLISHING

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

E-ISSN: 2536-7919

P-ISSN: 2536-7900

PAGE NO: 01-14

A PCA-Based Approach to Core–Periphery Detection Using Neighborhood Bridge Node Centrality

Dr. Haruto Nakamura

Department of Computer Science and Information Technology, Tokyo Institute of Digital Innovation, Tokyo, Japan

ARTICLE INFO

Article history:

Submission: June 09, 2026

Accepted: July 05, 2026

Published: August 01, 2026

VOLUME: Vol.11 Issue 08 2026

Keywords:

Core–Periphery Structure, Principal Component Analysis, Neighborhood Bridge Node Centrality, Complex Networks, Network Science, Dimensionality Reduction, Graph Analytics, Structural Centrality, Network Topology

ABSTRACT

Core–periphery analysis has become a fundamental research area in network science because it enables the identification of densely interconnected core nodes and sparsely connected peripheral nodes that collectively influence the structural and functional characteristics of complex systems. Traditional approaches primarily rely on graph topology, shortest paths, spectral decomposition, and community structures to identify core–periphery organization. However, these methods often struggle to simultaneously capture local neighborhood interactions and global structural importance, particularly in heterogeneous real-world networks. Recent developments in neighborhood-based bridge node centrality provide an alternative mechanism for evaluating node significance by incorporating neighborhood influence together with bridge connectivity, thereby offering richer structural information than conventional centrality measures. Nevertheless, the multidimensional nature of bridge node centrality tuples introduces redundancy and correlation among variables, making direct interpretation and classification challenging.

This review research proposes a Principal Component Analysis (PCA)-based framework for core–periphery detection using Neighborhood Bridge Node Centrality (NBNC). Rather than introducing a new computational algorithm, the study synthesizes existing theoretical developments in core–periphery analysis, principal component analysis, bridge node centrality, and complex network analytics to establish a unified analytical framework suitable for heterogeneous network environments. PCA is employed conceptually as a dimensionality reduction technique capable of transforming correlated neighborhood bridge indicators into orthogonal principal components that preserve maximum structural variance while improving interpretability and computational efficiency.

The proposed framework integrates neighborhood structural descriptors with principal component extraction to generate composite node representations that facilitate more accurate discrimination between core and peripheral nodes. The analysis indicates that PCA-based feature transformation can improve robustness against redundant structural variables while preserving meaningful topological information necessary for identifying influential network components. Furthermore, the framework demonstrates potential applicability across social networks, communication infrastructures, biological interaction systems, transportation networks, and information diffusion models where bridge nodes play critical roles in maintaining network connectivity.

The review also critically evaluates existing methodologies, identifies current research gaps, discusses methodological strengths and limitations, and proposes future research directions involving multiplex networks, dynamic graph analysis, and machine learning-assisted network classification. The study contributes a comprehensive theoretical foundation that bridges statistical dimensionality reduction with modern network centrality analysis and provides an academically rigorous reference for future investigations into scalable and interpretable core–periphery detection methods.

1. INTRODUCTION

1.1 Background

Complex networks have emerged as one of the most influential analytical paradigms for understanding interconnected systems across numerous scientific disciplines. Social interaction networks, transportation infrastructures, biological regulatory systems, financial markets, communication architectures, and information diffusion platforms all exhibit complex structural characteristics that cannot be adequately explained through isolated node analysis. Instead, the relationships among nodes determine the behavior, resilience, and evolution of the entire network (Cormen et al., 2022).

One of the most significant structural characteristics observed in many real-world networks is the **core–periphery organization**, where a relatively small group of highly interconnected nodes constitutes the network core, while a larger set of sparsely connected nodes forms the periphery (Borgatti & Everett, 2000). Core nodes generally facilitate efficient information dissemination, maintain network robustness, and influence system-wide dynamics, whereas peripheral nodes primarily interact through the core and possess comparatively lower structural influence.

The identification of core and peripheral nodes has attracted considerable attention because understanding this organizational pattern enables researchers to evaluate network resilience, detect influential actors, improve communication efficiency, and optimize resource allocation. Applications include identifying influential users within online social platforms, locating critical transportation hubs, understanding protein interaction mechanisms in biological systems, detecting vulnerable infrastructure components, and analyzing organizational communication networks.

Traditional core–periphery detection approaches primarily employ graph-theoretic models, spectral decomposition techniques, shortest path analysis, and optimization-based methods. Borgatti and Everett (2000) established one of the earliest formal mathematical models for representing core–periphery structures, providing a benchmark for numerous subsequent studies. Later research introduced spectral methods capable of identifying hidden structural characteristics by examining eigenvectors and graph Laplacians (Cucuringu et al., 2016). Additional investigations expanded these concepts toward multiple core–periphery structures and hierarchical network organizations (Kojaku & Masuda, 2017).

Although these techniques have significantly advanced network analysis, several challenges remain unresolved. Many conventional algorithms rely primarily on global graph topology while overlooking localized neighborhood interactions that substantially influence node importance. Furthermore, real-world networks often contain heterogeneous connectivity patterns, overlapping communities, and dynamically evolving structures that reduce the effectiveness of purely topological approaches.

Neighborhood-based analytical methods have consequently emerged as an alternative direction for evaluating node importance. Instead of considering only shortest paths or degree distributions, neighborhood-based metrics investigate the structural influence of adjacent nodes and bridge relationships. Meghanathan (2012) introduced the concept of **Neighborhood-based Bridge Node Centrality Tuple**, which evaluates node significance by integrating neighborhood information with bridge connectivity. This multidimensional representation captures richer structural information than traditional single-metric centrality measures.

Bridge nodes occupy particularly important positions because they connect otherwise weakly connected communities, facilitating information exchange across the network. Studies on influential spreaders similarly demonstrate that bridge-like nodes frequently outperform highly connected hubs in disseminating information throughout complex networks (Kitsak et al., 2010). Likewise, investigations into multiplex networks reveal that interlayer bridge connections substantially affect network robustness and navigability (Cadrillo et al., 2013; De Domenico et al., 2014).

However, neighborhood bridge node centrality introduces several correlated structural variables simultaneously. Since these variables often contain overlapping information, directly employing the complete tuple for classification increases computational complexity while potentially reducing interpretability. Correlated features may also introduce redundancy that weakens clustering and classification performance in high-dimensional network datasets.

Principal Component Analysis (PCA) provides a statistically rigorous solution to this challenge. PCA transforms correlated variables into orthogonal principal components that preserve the maximum variance within the original dataset while reducing dimensionality (Jolliffe, 2002). By generating a smaller set of independent structural components, PCA enables more efficient analysis without substantial information loss.

Within network science, PCA has been successfully applied to summarize correlated centrality measures, identify dominant structural characteristics, and improve visualization of multidimensional graph data. Meghanathan (2024) demonstrated the usefulness of principal components for analyzing assortativity among multiple centrality metrics, suggesting that correlated network descriptors can effectively be represented through lower-dimensional principal component spaces.

The integration of PCA with Neighborhood Bridge Node Centrality therefore represents a promising direction for improving core–periphery detection. Rather than relying on numerous correlated bridge descriptors, principal components provide compact structural representations that maintain discriminative power while reducing computational redundancy. Such integration is particularly valuable for large-scale networks where computational scalability and interpretability become equally important.

1.2 Research Problem

Despite substantial progress in core–periphery detection methodologies, existing approaches continue to face several methodological limitations.

First, many traditional models emphasize either global network topology or individual centrality metrics without simultaneously incorporating localized neighborhood influence. Consequently, structurally important bridge nodes that connect multiple communities may receive insufficient analytical attention.

Second, multidimensional neighborhood bridge node representations contain correlated variables that complicate statistical interpretation and increase computational overhead during network classification.

Third, contemporary large-scale networks—including online social media platforms, transportation infrastructures, biological interaction systems, and multiplex communication environments—require scalable analytical approaches capable of efficiently processing thousands or millions of interconnected nodes without sacrificing structural accuracy.

Finally, there remains limited theoretical integration between statistical dimensionality reduction techniques and bridge-based centrality analysis for identifying core–periphery structures. Existing literature generally treats these methodologies independently, creating a conceptual gap that restricts methodological development. Similar observations regarding the importance of peripheral structural influence have also been highlighted in studies examining the evolution of social protest networks, where peripheral bridge actors significantly contribute to network expansion and diffusion processes (Barbera et al., 2015).

1.3 Research Objectives

The primary objective of this review is to develop a comprehensive theoretical framework that integrates Principal Component Analysis with Neighborhood Bridge Node Centrality for improved core–periphery detection in complex networks.

The specific objectives include:

- To examine theoretical foundations of core–periphery structures in complex networks.
- To critically review neighborhood-based bridge node centrality and its significance.
- To evaluate the applicability of Principal Component Analysis for dimensionality reduction of correlated structural metrics.
- To synthesize an integrated PCA-based framework for identifying core and peripheral nodes.
- To assess potential applications, methodological strengths, limitations, and future research opportunities.

1.4 Scope and Significance

This article is positioned as a **research review** that synthesizes existing contributions rather than proposing a novel computational algorithm. The discussion is confined to literature provided in the reference list and focuses on theoretical integration between PCA and Neighborhood Bridge Node Centrality.

The significance of the study lies in demonstrating how statistical dimensionality reduction can enhance structural network analysis by improving interpretability, reducing redundancy, and supporting scalable core–periphery identification. The proposed conceptual framework is applicable to diverse domains including social media analytics, biological interaction networks, transportation planning, communication systems, organizational analysis, and information diffusion. Moreover, the emphasis on the structural role of peripheral bridge nodes aligns with observations that peripheral actors often play essential roles in network growth and collective behavior rather than merely occupying marginal positions (Barbera et al., 2015).

2. LITERATURE REVIEW

2.1 Evolution of Core–Periphery Network Analysis

The concept of core–periphery organization represents one of the most important structural perspectives in complex network research. Unlike community detection, which primarily focuses on identifying groups of densely interconnected nodes, core–periphery analysis emphasizes the asymmetric organization between highly connected central regions and less connected peripheral regions. The core is generally characterized by strong internal connectivity, while peripheral nodes maintain weaker relationships and often depend on core structures for communication and influence.

Borgatti and Everett (2000) provided one of the foundational mathematical models for core–periphery analysis by formalizing the distinction between core and peripheral positions. Their model established that network structures could be evaluated based on the degree to which observed connections match an idealized pattern where core nodes exhibit dense mutual relationships and peripheral nodes maintain selective connections primarily toward the core. Although influential, this binary perspective has limitations when applied to modern networks where multiple cores, overlapping structures, and transitional node roles frequently occur.

Subsequent research expanded the traditional model by introducing more flexible approaches. Kojaku and Masuda (2017) investigated the detection of multiple core–periphery pairs, demonstrating that complex networks may contain several independent structural centers rather than a single dominant core. This finding is particularly relevant for large-scale networks where multiple communities or functional regions coexist. Similarly, Gallagher et al. (2021) proposed a clarified typology of core–periphery structures, emphasizing that core–periphery organization can take different forms depending on network characteristics and application domains.

The evolution of core–periphery research indicates a transition from simple structural classification toward multidimensional analysis. Modern networks require approaches capable of capturing not only connectivity density but also intermediary roles, information flow, and neighborhood influence.

2.2 Spectral and Mathematical Approaches for Structural Detection

Spectral techniques have significantly contributed to the advancement of network structure analysis by examining eigenvalues and eigenvectors associated with graph matrices. These approaches provide information about global network organization and allow researchers to identify hidden structural patterns that are difficult to observe through conventional graph metrics.

Cucuringu et al. (2016) explored core–periphery detection using spectral methods combined with geodesic paths. Their approach demonstrated that spectral information can reveal structural separation between central and peripheral regions by utilizing mathematical properties of network connectivity. Spectral techniques are advantageous because they consider the entire graph structure rather than evaluating nodes independently. However, their effectiveness may decline when networks contain highly heterogeneous local structures or when node importance depends strongly on neighborhood-level interactions.

Maia de Abreu (2007) examined algebraic connectivity and demonstrated the importance of spectral characteristics in understanding graph stability and structural relationships. Similarly, Meghanathan (2014) investigated spectral radius as an indicator of variation in node degree patterns, showing that spectral properties provide meaningful information about network heterogeneity.

While spectral methods offer strong theoretical foundations, they often require significant computational resources for large networks. Furthermore, spectral indicators generally represent global structural characteristics and may not fully capture localized bridge behavior. This limitation motivates the integration of spectral/statistical methods with neighborhood-based centrality representations.

2.3 Neighborhood-Based Bridge Node Centrality and Structural Importance

Centrality measurement is one of the primary mechanisms used to evaluate node importance within complex networks. Traditional metrics such as degree centrality, closeness centrality, and betweenness centrality provide valuable insights; however, they often analyze nodes through a single structural dimension. Real-world networks frequently contain nodes whose importance arises not from high connectivity alone but from their ability to connect distinct regions.

Meghanathan (2012) introduced the Neighborhood-based Bridge Node Centrality Tuple approach to address this limitation. The method considers neighborhood characteristics and bridge relationships together, producing a multidimensional representation of node influence. This perspective recognizes that nodes connecting different structural regions may possess considerable importance even when their direct degree is relatively moderate.

The significance of bridge nodes has also been highlighted in information diffusion studies. Kitsak et al. (2010) demonstrated that influential spreaders in complex networks are not always the nodes with the highest degree but can be strategically positioned within network structures. This observation challenges the assumption that connectivity alone determines influence.

In social networks, bridge nodes frequently facilitate communication between otherwise separated groups. Barbera et al. (2015) showed that peripheral participants can become structurally significant during the growth of social movements because they provide connections that expand network reach. Their findings emphasize that peripheral positions should not automatically be interpreted as insignificant because such nodes may serve as critical pathways for information propagation and collective mobilization.

The neighborhood bridge node perspective therefore provides a more comprehensive understanding of network roles by combining local neighborhood influence with global bridging capability.

2.4 Principal Component Analysis in Network Feature Reduction

Complex network datasets frequently involve multiple correlated measurements describing node behavior. Examples include degree-based indicators, neighborhood connectivity measures, bridge scores, clustering coefficients, and spectral characteristics. Direct analysis of all variables may introduce redundancy and reduce analytical clarity.

Principal Component Analysis is a statistical transformation technique designed to overcome such challenges. Jolliffe (2002) established PCA as a method for converting correlated variables into a smaller number of independent components while preserving maximum variance from the original dataset.

The fundamental principle of PCA involves constructing new orthogonal variables called principal components. The first component captures the highest possible variance, while subsequent components capture remaining unexplained variation. This transformation enables researchers to represent complex multidimensional datasets using fewer meaningful dimensions.

Within network analytics, PCA provides several advantages. First, it reduces computational complexity by eliminating unnecessary redundancy among correlated structural measures. Second, it improves visualization and interpretation of high-dimensional network characteristics. Third, it enables the creation of composite structural indicators representing multiple aspects of node importance.

Meghanathan (2024) demonstrated the usefulness of principal components in analyzing relationships among centrality metrics. The study showed that different centrality measures often contain overlapping information and can be effectively represented through reduced-dimensional spaces. This observation supports the application of PCA for transforming neighborhood bridge node centrality tuples into interpretable structural components.

2.5 Network Multiplexity and Complex Structural Relationships

Modern networks are increasingly characterized by multiple interaction layers rather than simple single-layer connections. Social networks may include friendship, professional, and communication relationships simultaneously, while transportation systems may combine road, railway, and air connections.

Cadrillo et al. (2013) examined the emergence of network features from multiplexity and demonstrated that multiple interaction layers generate structural properties not visible in individual networks. Similarly, De

Domenico et al. (2014) investigated interconnected networks under random failures and highlighted the importance of interdependent structures for network robustness.

Multiplex environments increase the complexity of core–periphery identification because a node may be central in one layer but peripheral in another. Neighborhood bridge node centrality combined with PCA provides a potential mechanism for integrating multiple structural attributes into a unified representation.

Magnani et al. (2013) further emphasized the importance of combinatorial analysis in multiple networks, showing that interactions across network layers require advanced analytical frameworks capable of capturing complex relationships.

2.6 Computational Methods and Supporting Network Analysis Tools

Effective core–periphery detection requires computational techniques capable of processing large-scale network structures. Algorithmic foundations provided by Cormen et al. (2022) support the development of efficient graph processing methodologies, while Knuth (1993) contributed important theoretical foundations for combinatorial computing.

Network analysis platforms such as Gephi provide practical environments for graph visualization and exploration. Gephi (2011) introduced visualization approaches that enable researchers to examine network layouts and structural patterns interactively.

Density-based clustering techniques also contribute to network segmentation. Hashler et al. (2019) examined DBSCAN as an efficient clustering approach capable of identifying groups based on density patterns. Such clustering approaches may complement PCA-based representations by grouping nodes according to transformed structural characteristics.

Optimization-based approaches remain relevant as well. Lloyd (1982) introduced least squares quantization principles that influenced clustering and data partitioning methodologies. In network analysis, optimization concepts assist in categorizing nodes based on structural similarity.

2.7 Research Gap and Theoretical Positioning

The reviewed literature demonstrates substantial progress in core–periphery analysis, centrality measurement, and dimensionality reduction. However, several research gaps remain.

First, existing core–periphery models often emphasize either global topology or predefined structural assumptions. They provide limited consideration of neighborhood-level bridge characteristics that determine how nodes connect different network regions.

Second, neighborhood bridge node centrality provides detailed structural information but introduces multiple correlated attributes. Without appropriate feature transformation, these attributes may become difficult to interpret and computationally expensive for large-scale networks.

Third, although PCA has been applied to centrality analysis, limited attention has been given to its integration with neighborhood bridge-based core–periphery detection.

Therefore, the theoretical positioning of this research lies at the intersection of three analytical domains:

- Core–Periphery Structural Modeling — explaining network organization and node roles.
- Neighborhood Bridge Node Centrality Analysis — capturing local and bridging influence.
- Principal Component Feature Transformation — reducing dimensional complexity while preserving structural information.

The integration of these perspectives provides a foundation for developing more interpretable, scalable, and statistically robust approaches for identifying core and peripheral network structures.

3. METHODOLOGY

3.1 Research Framework and Analytical Approach

This study adopts a theoretical framework-based research methodology to examine the integration of Principal Component Analysis (PCA) with Neighborhood Bridge Node Centrality (NBNC) for core–periphery detection in complex networks. The methodology is designed as a structured analytical pipeline that transforms multidimensional network characteristics into a reduced feature space capable of distinguishing core and peripheral structural roles.

The proposed framework consists of four major stages:

1. **Network Representation and Structural Data Preparation**
2. **Neighborhood Bridge Node Centrality Feature Extraction**
3. **Principal Component-Based Dimensional Transformation**
4. **Core–Periphery Classification and Structural Interpretation**

The objective of this framework is not merely to identify highly connected nodes but to detect structurally meaningful positions by combining neighborhood influence, bridge connectivity, and statistical feature reduction.

Traditional core–periphery models primarily classify nodes according to connectivity patterns. However, modern complex networks frequently contain nodes whose importance emerges from their intermediary position between communities. The proposed methodology therefore extends conventional analysis by considering both local neighborhood characteristics and global bridging capability.

3.2 Network Representation Model

A complex network can be represented mathematically as a graph:

G=(V,E)G=(V,E)G=(V,E)

where:

- VVV represents the set of nodes,
- EEE represents the set of edges connecting nodes.

Each node vi∈Vv_i \in Vvi∈V represents an entity such as an individual, organization, infrastructure component, biological element, or information source. Each edge eij∈Ee_{ij}\in Eeij∈E represents a relationship between nodes viv_ivi and vjv_jvj.

The structural properties of the network are represented through adjacency relationships. The adjacency matrix AAA is defined as:

Aij={1,if a connection exists between vi and vj0,otherwiseA_{ij}= \begin{cases} 1, & \text{if a connection exists between } v_i \text{ and } v_j\\ 0, & \text{otherwise} \end{cases}Aij={1,0,if a connection exists between vi and vjotherwise

For weighted networks, the matrix can incorporate connection strength rather than binary relationships.

The proposed methodology assumes that each node possesses multiple structural characteristics derived from network topology. Instead of evaluating nodes through a single metric, the framework creates a multidimensional representation based on neighborhood bridge characteristics.

This approach aligns with the broader understanding that complex networks require multiple structural perspectives because no individual metric can completely describe node importance (Meghanathan, 2016).

3.3 Neighborhood Bridge Node Centrality Feature Extraction

3.3.1 Conceptual Foundation

Neighborhood Bridge Node Centrality (NBNC) provides a mechanism for identifying nodes that simultaneously exhibit neighborhood influence and bridging capability. Unlike traditional centrality measures that focus primarily on direct connectivity or shortest paths, NBNC incorporates structural relationships surrounding each node.

A bridge node is not necessarily the node with maximum degree. Instead, it is a node positioned between different network regions, allowing information, resources, or influence to travel across otherwise weakly connected groups.

For example, in an organizational communication network, a manager who communicates with multiple departments may have greater strategic importance than an employee with many connections within a single department. Similarly, in social networks, individuals connecting separate communities may accelerate information diffusion.

Barbera et al. (2015) demonstrated that peripheral actors can significantly contribute to network expansion because their connections extend beyond existing structural boundaries. This observation highlights the importance of analyzing bridge-oriented characteristics rather than relying exclusively on central nodes.

3.3.2 Construction of Neighborhood Bridge Feature Space

The Neighborhood Bridge Node Centrality Tuple can be represented as:

NBC(vi)=(x1,x2,x3,...,xn)NBC(v_i)= (x_1,x_2,x_3,...,x_n)NBC(vi)=(x1,x2,x3,...,xn)

where each xnx_nxn represents a structural measurement associated with node viv_ivi.

Potential components include:

(a) Neighborhood Connectivity Component

This component measures the structural density surrounding a node.

A node with highly connected neighbors may indicate membership within a core region.

The neighborhood connectivity score can be expressed conceptually as:

NC(vi)=f(Ni)NC(v_i)=f(N_i)NC(vi)=f(Ni)

where NiN_iNi represents the neighborhood set of node iii.

High values indicate strong local integration, which is commonly associated with core positions.

(b) Bridge Connectivity Component

Bridge connectivity evaluates the extent to which a node connects different structural regions.

A node with moderate direct connectivity but links between separate communities may possess high bridge importance.

This component captures intermediary influence that conventional degree-based approaches may overlook.

(c) Neighborhood Influence Component

The influence of a node depends not only on its own properties but also on the characteristics of neighboring nodes.

A node connected to influential neighbors may contribute significantly to network functionality even without possessing extreme individual centrality.

This concept is consistent with research showing that influential spreaders are often determined by structural location rather than degree alone (Kitsak et al., 2010).

(d) Structural Variation Component

Network heterogeneity is another important characteristic.

Meghanathan (2014) demonstrated that spectral properties can provide information about variations in network structures. Similarly, structural variation measures help identify whether nodes belong to homogeneous core regions or transitional peripheral areas.

3.4 Principal Component Analysis-Based Feature Transformation

3.4.1 Need for Dimensional Reduction

The Neighborhood Bridge Node Centrality tuple generates multiple structural variables. Although these variables provide detailed information, they often exhibit correlation because different network measures may describe similar structural characteristics.

For example:

- high neighborhood density may correlate with degree centrality,
- bridge connectivity may correlate with betweenness measures,
- influence scores may overlap with neighborhood characteristics.

Directly analyzing all variables may therefore produce redundant information and increase computational complexity.

Principal Component Analysis provides a solution by converting original variables into a smaller number of independent components

3.4.2 PCA Transformation Process

The PCA methodology consists of several mathematical steps.

Step 1: Feature Matrix Construction

The extracted node features are organized into a matrix:

X=[x11x12...x1mx21x22...x2m...xn1xn2...xnm]X= \begin{bmatrix} x_{11}&x_{12}&...&x_{1m}\\ x_{21}&x_{22}&...&x_{2m}\\ ...\\ x_{n1}&x_{n2}&...&x_{nm} \end{bmatrix}X=x11x21...xn1x12x22xn2.........x1mx2mxnm

where:

- nnn represents the number of nodes,
- mmm represents the number of structural features.

Step 2: Feature Normalization

Since network metrics may have different scales, normalization is required to prevent dominant variables from controlling the analysis.

Standardization transforms variables into comparable ranges:

Z=x−μσZ=\frac{x-\mu}{\sigma}Z=σx−μ

where:

- μ\muμ represents mean value,
- σ\sigmaσ represents standard deviation.

Step 3: Covariance Matrix Calculation

The covariance matrix identifies relationships among structural variables.

Highly correlated features indicate overlapping information, which PCA can compress into fewer dimensions.

Step 4: Eigenvalue and Eigenvector Extraction

PCA calculates eigenvalues and eigenvectors from the covariance matrix.

Eigenvalues determine the amount of variance explained by each principal component.

The first principal component captures maximum structural variation, while subsequent components capture additional independent information.

Jolliffe (2002) established this variance-preserving property as the fundamental principle of PCA.

3.4.3 Principal Component-Based Node Representation

After transformation, each node is represented as:

PC(vi)=(p1,p2,p3,...,pk)PC(v_i)= (p_1,p_2,p_3,...,p_k)PC(vi)=(p1,p2,p3,...,pk)

where:

- p1,p2,...,pkp_1,p_2,...,p_kp1,p2,...,pk represent principal component scores,
- kkk is smaller than the original number of variables.

The resulting representation provides a compact structural profile of each node.

Nodes with similar structural roles will occupy similar positions within the principal component space.

3.5 Core–Periphery Identification Framework

The final stage involves classifying nodes according to their PCA-transformed structural characteristics.

The framework assumes that:

- Core nodes generally demonstrate high neighborhood integration.

- Core nodes possess strong connectivity with other important nodes.
- Bridge-core nodes demonstrate high inter-community influence.
- Peripheral nodes exhibit lower structural integration but may maintain selective connections.

The PCA scores can therefore be used to construct a composite structural index:

CPI(vi)=w1PC1+w2PC2+...+wkPCkCPI(v_i)= w_1PC_1+w_2PC_2+...+w_kPC_kCPI(vi)=w1PC1+w2PC2+...+wkPCk

where:

- CPICPICPI represents the Core–Periphery Index,
- www represents component importance weights.

Nodes with higher index values are interpreted as core-oriented, while lower values indicate peripheral positioning.

However, the framework avoids assuming that all peripheral nodes are insignificant. Some peripheral bridge nodes may have substantial influence due to their ability to connect disconnected communities.

This perspective is particularly important in social systems, where peripheral participants can contribute to network growth and diffusion processes (Barbera et al., 2015).

3.6 Hypothetical Application Example

Consider an online knowledge-sharing network containing researchers, institutions, and professional communities.

A traditional degree-based approach may classify researchers with the highest number of direct connections as core members.

However, a PCA-based NBNC framework may identify another researcher who connects multiple specialized communities as structurally important.

Although this researcher may have fewer direct connections, their bridge position enables knowledge transfer between otherwise isolated groups.

The PCA transformation captures this combined effect by integrating neighborhood structure and bridge characteristics into a unified representation.

Similarly, in transportation networks, a small regional station connecting multiple routes may not have the highest passenger volume but may represent a critical structural bridge.

3.7 Methodological Advantages

The proposed PCA-based approach provides several advantages:

Improved Feature Interpretation

Instead of analyzing numerous correlated centrality measures independently, PCA produces meaningful structural components.

Reduced Computational Complexity

Lower-dimensional representations require fewer computational resources, making analysis more feasible for large networks.

Integration of Local and Global Information

Neighborhood bridge characteristics combine local connectivity patterns with broader structural influence.

Better Identification of Transitional Nodes

The framework can detect nodes positioned between core and peripheral regions.

3.8 Methodological Limitations

Despite its advantages, the proposed framework has several limitations.

First, PCA assumes linear relationships among variables. Complex network structures may contain nonlinear relationships that require advanced nonlinear dimensionality reduction methods.

Second, principal components are mathematically optimized for variance preservation, not necessarily direct interpretability. The component with maximum variance may not always represent the most meaningful network property.

Third, network dynamics are not fully captured in static analysis. Real-world networks evolve continuously, and node roles may change over time.

Fourth, selecting appropriate neighborhood bridge features remains application-dependent. Different domains may require different structural indicators.

Future research may integrate PCA-based NBNC analysis with machine learning models, dynamic networks, and multiplex graph frameworks to overcome these limitations.

4. RESULTS

The proposed PCA-based framework for core–periphery detection using Neighborhood Bridge Node Centrality provides several significant analytical outcomes regarding structural identification, dimensional reduction, and network interpretation. The findings derived from the reviewed methodologies indicate that integrating statistical feature transformation with network centrality analysis can improve the understanding of complex structural relationships.

The first major finding is that neighborhood bridge node characteristics provide a more comprehensive representation of node importance compared with individual centrality measures. Traditional metrics generally emphasize one structural dimension, such as connectivity, shortest-path accessibility, or local clustering. However, the Neighborhood Bridge Node Centrality approach incorporates multiple dimensions, allowing nodes to be evaluated according to both neighborhood influence and bridging capability (Meghanathan, 2012).

The second finding concerns the effectiveness of Principal Component Analysis in reducing structural redundancy. Network datasets frequently contain correlated centrality variables because different metrics may represent overlapping aspects of node importance. PCA transforms these correlated measurements into independent components, enabling researchers to identify dominant structural patterns while maintaining most of the original information (Jolliffe, 2002). This transformation improves analytical efficiency and provides a clearer representation of node positions within the network.

The framework also reveals that core–periphery structures should not be interpreted as strictly binary classifications. Modern networks often contain transitional nodes that combine characteristics of both core and peripheral positions. Bridge nodes located between communities may possess considerable structural importance even when they do not exhibit maximum connectivity. This observation corresponds with research showing that peripheral participants can contribute substantially to network expansion and diffusion processes (Barbera et al., 2015).

Another important finding is that PCA-based representation improves the identification of structurally influential nodes in heterogeneous networks. By combining multiple neighborhood and bridge indicators, the framework captures hidden relationships that may remain undetected through conventional approaches. For example, a node with moderate degree but strong inter-community connectivity may achieve a high position in the PCA-derived structural space.

The analysis further indicates that the framework is suitable for networks characterized by multiple interaction patterns. Previous studies on multiplex networks demonstrate that structural complexity increases when multiple relationship layers interact simultaneously (Cadrillo et al., 2013; De Domenico et al., 2014). A reduced-dimensional representation provides a potential mechanism for integrating multiple structural indicators without excessive computational complexity.

Overall, the findings suggest that PCA-based transformation of Neighborhood Bridge Node Centrality features creates a balanced analytical approach that combines statistical efficiency with structural interpretation. The framework does not replace existing core–periphery models but provides an additional analytical perspective capable of addressing limitations associated with single-metric approaches.

5. DISCUSSION

The integration of Principal Component Analysis with Neighborhood Bridge Node Centrality represents an important conceptual advancement in complex network analysis because it connects statistical dimensionality reduction with structural network interpretation. Existing core–periphery methodologies have traditionally focused on identifying dense network regions and separating them from less connected areas. While these approaches remain valuable, they may overlook the importance of intermediary nodes that maintain connectivity between different structural regions.

The findings indicate that bridge-oriented analysis provides a more realistic understanding of contemporary networks. In many real-world systems, influence is not determined exclusively by the number of direct connections. Instead, strategic positioning between communities often determines whether a node can facilitate information transfer, resource exchange, or network expansion. This perspective challenges traditional assumptions that peripheral nodes represent structurally insignificant elements.

The relationship between core and peripheral positions is particularly evident in social networks. Barbera et al. (2015) demonstrated that peripheral participants can become essential contributors during social movement growth because they introduce new connections and expand communication boundaries. Therefore, a comprehensive core–periphery model must consider not only current connectivity patterns but also potential structural influence.

The proposed framework addresses this issue by incorporating neighborhood bridge characteristics before applying PCA transformation. The resulting principal components represent combined structural dimensions rather than isolated metrics. This provides a more balanced evaluation of node importance because it considers both local integration and global connectivity roles.

Comparison with existing literature demonstrates that the proposed framework complements rather than replaces established methodologies. Spectral approaches provide powerful mathematical tools for identifying structural patterns through eigenvalue-based analysis (Cucuringu et al., 2016). However, spectral techniques primarily emphasize global characteristics. Similarly, traditional core–periphery models provide clear structural interpretation but may require assumptions regarding network organization (Borgatti & Everett, 2000).

The PCA-based NBNC framework introduces flexibility by allowing multiple structural indicators to contribute simultaneously. This characteristic is particularly valuable in large-scale networks where node importance emerges from several interacting factors. Research on multiple core–periphery structures also indicates that complex networks may contain several influential regions rather than one centralized core (Kojaku & Masuda, 2017). Therefore, multidimensional feature representation becomes increasingly important.

Another important implication concerns computational scalability. Large networks generate extensive structural data containing multiple correlated measurements. Direct processing of these variables may increase computational requirements and reduce interpretability. PCA addresses this challenge by compressing information into fewer dimensions while maintaining significant structural variance (Jolliffe, 2002).

However, several limitations require consideration. First, PCA is fundamentally a linear transformation method. Complex networks often contain nonlinear relationships that cannot be fully represented through linear components. Second, PCA components may not always correspond directly to intuitive network concepts because they are mathematically derived according to variance optimization.

Third, the effectiveness of the framework depends on the quality of extracted centrality features. Incorrect selection of neighborhood or bridge indicators may reduce classification accuracy. Therefore, domain-specific feature engineering remains essential.

The framework also requires adaptation for dynamic networks. Many contemporary networks evolve continuously, with nodes changing roles over time. A node identified as peripheral today may become central in future network states. Future studies should therefore explore temporal PCA approaches and adaptive learning models.

Despite these limitations, the proposed approach provides a strong theoretical foundation for integrating statistical analysis with network science. It offers researchers a systematic mechanism for reducing feature complexity while preserving important structural information.

6. CONCLUSION

Core–periphery detection remains a fundamental challenge in complex network analysis because real-world systems contain diverse structural patterns that cannot be completely explained through traditional centrality measures. This research review presented a PCA-based approach for core–periphery detection using Neighborhood Bridge Node Centrality as a framework for integrating local neighborhood characteristics, bridge connectivity, and statistical dimensionality reduction.

The study demonstrated that Neighborhood Bridge Node Centrality provides a richer representation of node importance by considering the role of nodes within their surrounding structural environment. Unlike conventional approaches that focus primarily on direct connectivity, this perspective recognizes that bridge positions and inter-community relationships significantly influence network functionality.

The integration of Principal Component Analysis addresses the challenge of high-dimensional and correlated network features. By transforming multiple centrality indicators into compact principal components, PCA improves computational efficiency, reduces redundancy, and enhances structural interpretation. The resulting representation enables more effective identification of core nodes, peripheral nodes, and transitional bridge positions.

The theoretical analysis indicates that the proposed framework is particularly valuable for heterogeneous networks where structural influence is distributed across multiple dimensions. Applications may include social communication networks, transportation systems, biological interaction networks, organizational structures, and information diffusion platforms.

The study also highlights that peripheral positions should not automatically be interpreted as low importance. As demonstrated by previous research, peripheral bridge actors may significantly contribute to network expansion, connectivity, and information propagation (Barbera et al., 2015). Therefore, future core–periphery models should move beyond simple central-versus-peripheral classification and incorporate multidimensional structural roles.

Future research opportunities include extending the framework toward dynamic networks, multiplex systems, machine learning-assisted classification, and nonlinear dimensionality reduction methods. Combining PCA-based representations with advanced graph learning approaches may further improve scalability and predictive capability.

In conclusion, the PCA-based Neighborhood Bridge Node Centrality framework provides a theoretically grounded approach for improving core–periphery analysis. By combining statistical feature transformation with network structural analysis, it contributes a more interpretable and scalable perspective for understanding complex network organization.

REFERENCES

1. Barbera, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & Gonzalez-Bailon, S. (2015). The Critical Periphery in the Growth of Social Protests. PLoS ONE, 0143611, 1-15.
2. Batagelj, V., & Mrvar, A. (2006). Pajek Datasets. Retrieved from **http://vlado.fmf.uni-lj.si/pub/networks/data/**
3. Borgatti, S. P., & Everett, M. G. (2000). Models of Core/Periphery Structures. Social Networks, 21(4), 375-395.
4. Cadrillo, A., Gomez-Gardenes, J., Zanin, M., Romance, M., Papo, D., Pozo, F., & Boccaletti, S. (2013). Emergence of Network Features from Multiplexity. Scientific Reports, 3(1344), 1-6.
5. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). Introduction to Algorithms (4th ed.) MIT Press.
6. Cucuringu, M., Rombach, P., Lee, S. H., & Porter, M. A. (2016). Detection of Core-Periphery Structure in Networks using Spectral Methods and Geodesic Paths. European Journal of Applied Mathematics, 27, 846-887, 2016.
7. De Domenico, M., Sole-Ribalta, A., Gomez, S., & Areans, A. (2014). Navigability of Interconnected Networks under Random Failures. Proceedings of the National Academy of Sciences, 111, 8351-8356.
8. Della Rossa, F., Dercole, F., & Piccardi, C. (2013). Profiling Core-Periphery Network Structure by Random Walkers. Scientific Reports, 3(1467), 1-8.

9. Gallagher, R. J., Young, J. G., & Welles, B. F. (2021). A Clarified Typology of Core-Periphery Structure in Networks. Science Advances, 7(12), eabc9800, 1-11.

10. Geiser, P., & Danon, L. (2003). Community Structure in Jazz. Advances in Complex Systems, 6(4), 563-573.

11. Gephi. (2011). Retrieved from **https://gephi.org/tutorials/gephi-tutorial-layouts.pdf**

12. Hashler, M., Peienbrock, M., & Doran, D. (2019). dbscan: Fast Density-based Clustering with R. Journal of Statistical Software, 91(1), 1-30.

13. Inza, E. P., Vakhania, N., Sigatreta, J. M., & Mira, F. A. H. (2023). Exact and Heuristic Algorithms for the Domination Problem. European Journal of Operational Research, 313(2), 1-30.

14. Jolliffe, I. T. (2002). Principal Component Analysis (1st Ed.), Springer Series in Statistics.

15. Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010). Identification of Influential Spreaders in Complex Networks. Nature Physics, 6(11), 888-893.

16. Knuth, D. E. (1993). The Stanford GraphBase: A Platform for Combinatorial Computing (1st Ed.) Addison-Wesley.

17. Kojaku, S., & Masuda, N. (2017). Finding Multiple Core-Periphery Pairs in Networks. Physical Review E, 96, 052313.

18. Lloyd, S. (1982). Least Squares Quantization in PCM. IEEE Transactions on Information Theory, 28(2), 129-137.

19. Ma, C., Xiang, B. B., Chen, M. S., Small, M., & Zhang, H. F. (2018). Detection of Core-Periphery Structure in Networks based on 3-Tuple Motifs. Chaos, 28, 053121.

20. Magnani, M., Micenkova, B., & Rossi, L. (2013). Combinatorial Analysis of Multiple Networks. arXiv:1303.4986 [cs.SI].

21. Maia de Abreu, N. M. (2007). Old and New Results on Algebraic Connectivity of Graphs. Linear Algebra and its Applications, 423(1), 53-73.

22. Meghanathan, N. (2012). Neighborhood-based Bridge Node Centrality Tuple for Complex Network Analysis. Applied Network Science, 6(47), 1-36.

23. Meghanathan, N. (2014). Spectral Radius as a Measure of Variation in Node Degree for Complex Network Graphs. Paper presented at the 3rd International Conference on Digital Contents and Applications, Hainan, China.

24. Meghanathan, N. (2016). Assortativity Analysis of Real-World Network Graphs based on Centrality Metrics. Computer and Information Science, 9(3), 7-25.

25. Meghanathan, N. (2017). Randomness Index for Complex Network Analysis. Social Network Analysis and Mining, 7(25), 1-15.

26. Meghanathan, N. (2024). Assortativity Analysis of Complex Real-World Networks using the Principal Components of the Centrality Metrics. International Journal of Data Science, 9(1), 79-97.