mathematics

MDPI

Article

SL-WLEN, a Novel Semi-Local Centrality Metric with Weighted Lexicographic Extended Neighborhood for Identifying Influential Nodes in Networks with Weighted Edges and Nodal Attributes

Maricela Fernanda Ormaza Morejón [1] and Rolando Ismael Yépez Moreira [2,*]

[1] School of Business and International Trade, Pontifical Catholic University of Ecuador Ibarra, Ibarra 100102, Ecuador; mformaza@pucesi.edu.ec
[2] School of Industrial Production, Cotacachi Higher University Institute, Cotacachi 100302, Ecuador
* Correspondence: iyepez@institutocotacachi.edu.ec

Abstract

The identification of influential nodes in complex networks modeling manufacturing environments is a critical aspect, especially when considering both structure and nodal attributes. This becomes particularly relevant given that conventional weighted centrality measures typically only consider edge weights while ignoring node heterogeneity. We present SL-WLEN (Semi-Local centrality with Weighted Lexicographic Extended Neighborhood), a novel centrality metric designed to overcome these limitations. Based on LRASP (Local Relative Average Shortest Path) and lexicographic ordering, SL-WLEN integrates topological structure and nodal attributes by combining local components (degree and nodal values). The incorporation of lexicographic ordering preserves the relative importance of nodes at each neighborhood level, ensuring that those with high values maintain their influence in the final metric without distortions from statistical aggregations. This method is applied and its robustness evaluated in a quality control network for chip manufacturing, comprising 1555 nodes representing critical process characteristics, with weighted connections indicating their degree of correlation. Finally, the metric was evaluated against other established methods using the SIR propagation model and Kendall's τ coefficient, demonstrating that SL-WLEN maintains consistent values across all analyzed test networks.

Keywords: semi-local centrality; identification of influential nodes; complex networks; lexicographic ordering; influential nodes; quality control

MSC: 05C82; 05C90

check for updates

Received: 1 March 2025
Revised: 25 April 2025
Accepted: 26 April 2025
Published: 15 August 2025

Citation: Ormaza Morejón, M.F.; Yépez Moreira, R.I. SL-WLEN, a Novel Semi-Local Centrality Metric with Weighted Lexicographic Extended Neighborhood for Identifying Influential Nodes in Networks with Weighted Edges and Nodal Attributes. *Mathematics* **2025**, *13*, 2614. https://doi.org/10.3390/math13162614

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Complex network analysis fundamentally focuses on studying its structure, dynamics, and interactions to understand the importance of its nodes and connections in information diffusion, resilience, and global information. This has garnered significant attention in recent years [1]. Processes such as synchronization, diffusion, and cascade effects are predominantly influenced by nodes with higher influence and connectivity [2]. Thus, the study of these problems holds theoretical relevance, which is reflected in its practical applications across fields including computational biology, computer science, social networks, and artificial intelligence [3].

According to the information provided by a network, centrality metrics are classified into three groups (local, semi-local, and global). Each evaluates node influence from the following two perspectives: their topological structure, which analyzes connections and positions within the network, and their valued nodal attributes, which consider quantifiable characteristics and weights assigned to each node to determine its importance in the network.

Local metrics are subdivided into the following two categories: those based on topological structure, such as degree centrality [4], which evaluates direct connections, PageRank [5], which analyzes nodes as web pages, and Trust-PageRank [6]; and those with nodal weights, such as node-weighted degree centrality [7], which incorporates weights through an $f(Wx)$ function, and the WNDegree/WNEDegree/WNEOpshalDegree variants [8] that integrate nodal attributes with topological structure.

Semi-local metrics can also be approached from the following two perspectives: those based on topological structure such as K-Shell [9], which identifies influential nodes through iterative removal, mixed degree decomposition [10], which considers residual degrees, semi-local centrality [11], which evaluates first and second-level neighbors, local structural centrality [12], and degree and importance of lines [13]; and those with nodal weights such as node-weighted harmonic centrality [14], which combines weights and geodesic distances, node-weighted betweenness centrality [15], which evaluates flows between nodes, modified node-weighted eigenvector centrality [16], and MCNDI [17], which integrates multiple indicators through the CRITIC method.

Global metrics represent the third group that utilizes information from the entire network. Among those based on classical structural topological approaches, notable examples include betweenness centrality [4], which analyzes shortest paths between nodes, closeness centrality [17], which evaluates proximity to other nodes [18], and eigenvector centrality [19], which considers the importance of neighboring nodes; while from the nodal attributes perspective, they incorporate developments such as LARSP [20] and LASP [21] that optimize shortest path calculations and ARP [22] that considers reciprocal distances in directed networks.

These metrics have attempted to provide a balance between accuracy and efficiency in complex network analysis; however, local and global metrics possess limitations. Local metrics exhibit constraints, as they only consider highly restricted information from nodes' immediate neighborhood [23]. While computationally simple and efficient in considering only the nearest neighbors, their capability to identify truly influential nodes is compromised by this limited network vision. Meanwhile, global metrics, although more accurate by utilizing information from the entire network, face considerable practical challenges. Their high computational complexity makes them impractical for large-scale networks [24].

Semi-local metrics, particularly those implementing the extended neighborhood concept (ENC), overcome these limitations by providing an optimal balance. By considering local subgraphs with LRASP (Local Relative Average Shortest Path), they generally achieve high accuracy in identifying influential nodes while maintaining manageable computational complexity [25]. This approach enables the evaluation of both topological position and semi-local structure, simultaneously considering node importance and the influence of its nearby neighbors.

Despite advances in semi-local metrics such as LASP, which incorporates LRASP and ENC to evaluate centrality by combining topological structure, there remains a significant gap in developing metrics that effectively integrate both topological structure and nodal values at the semi-local level. While metrics such as node-weighted harmonic centrality evaluate nodal weights with geodesic distances and those such as node-weighted betweenness centrality consider flows between weighted nodes, these analyze weights in isolation

without considering how these values affect the structure of local connections. This separation between weight and structure is particularly problematic in networks where a node's influence depends on both factors in an interrelated manner, as occurs in phenomena such as quality control in manufacturing environments or scientific collaboration networks, where both node attributes and network position jointly determine their actual importance. For instance, the works of [26–30] have addressed quality control through complex network analysis, although limiting themselves to the study of topological structure.

Lexicographic ordering has been utilized in various complex network contexts. Notable applications include the study of information diffusion through nodal configuration mapping [31] and node importance evaluation through minimal winning coalitions [32]. However, its potential for integrating topological structure with nodal values in centrality metrics remains relatively unexplored. This gap motivates the development of a new metric that leverages lexicographic ordering properties to simultaneously evaluate structure and nodal values in specific testing contexts, such as quality control.

In this context, this paper proposes SL-WLEN (Semi-Local Centrality with Weighted Lexicographic Extended Neighborhood), a novel centrality metric specifically designed for weighted networks with valued nodal attributes. The proposed method was applied, and its robustness was evaluated in a quality control network for semiconductor chip manufacturing, which was composed of 1555 nodes representing critical characteristics of the production process, with weighted connections indicating their degree of correlation between these variables. The metric was evaluated against other established methods in the scientific literature using the SIR propagation model and Kendall's τ coefficient, demonstrating that SL-WLEN maintained notably consistent values across all analyzed test networks, which validated its effectiveness for identifying influential nodes in complex manufacturing environments.

The methodology of this work encompasses the construction of a quality control process network for chip production followed by the theoretical formulation of the new SL-WLEN metric and its practical implementation, and it culminates with a robustness and efficacy analysis compared to other well-established models.

2. Establishment of a Quality Control Process Network for Chip Production

Complex network theory constitutes a viable methodology for analyzing and modeling interrelationships in quality control systems for chip manufacturing [29]. By establishing a network model that maps the evolution of critical quality parameters during the production process, it becomes possible to precisely identify crucial control points in the manufacturing chain. This approach enables the visualization of how each stage in the chip manufacturing process influences subsequent stages, facilitating the early detection of potential quality deviations [27].

To create a network that represents the dynamics of quality control in the chip manufacturing process, it is necessary to analyze and process information regarding specific characteristics that influence finished product quality. This enables the definition of each mode and how they relate and interact with one another.

2.1. Baseline Information Configuring the Network

The data used to configure the network consist of two sets. The first is a matrix $X \in R^{Nxp}$, where $N = 1763$ corresponds to the number of observations or manufactured

products, and $p = 1555$ represents the quality characteristics. The matrix is defined according to the following Equation (1):

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1j} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2j} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & x_{i3} & \dots & x_{ij} & \dots & x_{ip} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{nj} & \dots & x_{np} \end{bmatrix} \quad (1)$$

where $x_{ij} \in \{0,1\}, \forall i,j \in \mathbb{N}$ with $1 \leq \mathrm{i} \leq n, 1 \leq \mathrm{j} \leq p$.

Each row $x_i = \{x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{ip}\}$ represents the characteristics of the *i*-th product. Each column vector $x_i = \{x_{1j}, x_{2j}, \dots, x_{ij}, \dots, x_{nj}\}$ represents the values of the *j*-th characteristic for all products. Each component x_{ij} indicates the presence (1) or absence (0) of quality defects in the *j*-th characteristic for the *i*-th product. Each product is manufactured in the same system. The second set corresponds to a vector $Y \in R^N$, where $N = 1763$. Each element of vector Y contains information associated with the quality of each product or observation from matrix X. Specifically, it indicates whether the i-th finished product meets the required final quality (1) or is defective (0).

2.2. Network Node Definition

Within the network model, each node represents a quality characteristic of the manufactured product, and the nodal value of each node is defined based on a logistic regression model with Lasso $(L1)$ regularization. This allows for the assignment of a numerical value v_i to each node, representing the relevance or influence of the corresponding characteristic on the manufactured product's quality. The objective function for logistic regression with Lasso regularization is expressed according to the following Equation (2):

$$\begin{aligned} -\tfrac{1}{N}\sum_{i=1}^{N}\{y_i log P(Y=1|x_i) + (1-Y_i) log P(Y=0|x_i)\} + \lambda\|\beta\|_1 \\ = -\tfrac{1}{N}\sum_{i=1}^{N}\left\{y_i(\beta_0 + \beta^T x_i) - log\left(1 + e^{\beta_o + \beta^T x_i}\right)\right\} + \lambda\|\beta\|_1 \end{aligned} \quad (2)$$

where N is the total number of observations in the dataset, (y_i) is the i-th observation of the binary dependent variables, $X \in R^{Nxp}$ is the matrix of feature vectors (independent variables) for all N observations, and $Y \in R^N$ is the binary dependent vector, with β being the model coefficients and λ the regularization parameter.

The equation consists of the following two components: the first is the negative log-likelihood expression for binary logistic regression: $-\frac{1}{N}\sum_{i=1}^{N}\{y_i \log P(Y=1|x_i) + (1-Y_i) \log P(Y=0|x_i)\}$, obtained after applying the negative natural logarithm to the original likelihood function (Equation (2)), while the second component, the $L1$ penalty, integrates two elements, λ representing the $L1$ regularization parameter and the norm $\|\beta\|_1$.

Regarding the expression $\lambda\|\beta\|_1$, the L1 norm ($\|\beta\|_1$) of vector β is defined as the sum of the absolute values of its components. In other words, for vector $\beta = (\beta_1, \beta_2, ..., \beta_n)$, the norm is expressed as $\|\beta\|_1 = |\beta_1| +, |\beta_2|, + \cdots + |\beta_{n|}$, which measures the total magnitude of the coefficients. λ is a regularization parameter that controls the strength of the penalty. The larger λ is, the greater the penalty, leading to smaller coefficients. This additional penalty has the effect of 'shrinking' some coefficients towards zero and, in some cases, may cause certain coefficients to be exactly zero.

The incorporation of penalty $\lambda\|\beta\|_1$ into the objective function of the LASSO logistic regression enables automatic feature selection. By forcing some coefficients to zero, the

lasso tends to select a more relevant subset of features, eliminating less important ones. This could result in simpler and more generalizable models.

From the fitted model, the resulting β coefficients are utilized as nodal values v_i within the network model. Each node i represents a quality characteristic of the manufactured product, and its nodal value v_i is defined by the absolute magnitude of its estimated coefficient in the Lasso model (Equation (3)).

$$v_i = |\beta_i|, \forall_i \in \{1, 2, \ldots, p\} \tag{3}$$

2.3. Edge Weight Determination

The connections in the quality characteristics network are established through the Phi (φ) coefficient, which quantifies the degree and direction of the statistical association between pairs of binary characteristics in matrix $X \in R^{Nxp}$. For each pair of characteristics i, j, the φ coefficient defines the edge connecting them, evaluating the actual correlation between their variation patterns. This coefficient is calculated using the formula $\varphi_{ij} = \frac{(ad-bc)}{\sqrt{((a+b)+(c+d)+(a+c)+(b+d))}}$, where a, b, c, d correspond to the frequencies in the 2×2 contingency table between characteristics, i, j: a is the positive coincidence frequency (1, 1), b is the frequency of combination (1, 0), c is the frequency of combination (0, 1), and d is the negative coincidence frequency (0, 0).

The weight of each edge w_{ij} is defined through a threshold function applied to the φ coefficient. The function establishes that $w_{ij} = \varphi_{ij}$ if $\varphi_{ij} \geq \tau$, or $\varphi_{ij} \leq -\tau$, and it is equal to 0 if $-\tau < \varphi_{ij} < \tau$, where τ represents a statistically significant threshold. This threshold filters weak correlations, allowing only statistically significant relationships to form part of the network structure.

The topological structure of the graph is described by the adjacency matrix.

$A \in R^{pxp}$, where $A_{ij} = 1$ if nodes i, j are connected, and $A_{ij} = 0$ otherwise. The φ coefficient has a range of $[-1, 1]$, with extreme values indicating perfect association: $\varphi_{ij} = 1$ for perfect positive association, $\varphi_{ij} = -1$ for perfect negative association, and $\varphi_{ij} = 0$ for absence of association. This enables the construction of a network that faithfully reflects relationships between quality characteristics, capturing both positive and negative associations while avoiding irrelevant connections that could introduce noise into the analysis.

The symmetry of the φ coefficient $\varphi_{ij}(\varphi_{ij} = \varphi_{ji})$ and its specificity for binary variables make it ideal for modeling complex processes, such as chip manufacturing.

2.4. Construction of the Quality Control Network in Chip Manufacturing

During the chip manufacturing process, quality emerges as a complex phenomenon resulting from the dynamic interaction among multiple characteristics. This work adopts a complex network-based approach, visualizing quality control as an integrated system where each characteristic influences both individually and through its interactions with others.

The network is constructed by representing each quality characteristic as a node, whose importance is determined through the analysis of historical production data. Edges between nodes represent significant correlations between characteristics, revealing how changes in one can propagate and affect others. The resulting structure is an undirected weighted network, where nodal values quantify the individual importance of each characteristic, while edge weights reveal the strength of relationships between them. This model enables the visual understanding of how the production system's equilibrium depends on both individual characteristics and their complex network of interactions.

3. Definition of a Centrality Metric for Identification and Categorization of Quality Characteristics Based on the Network

3.1. Literature Review

The study of complex networks provides methodological frameworks and fundamental structures that enable the development of more advanced and sophisticated artificial intelligence systems [33]. The intersection between AI and complex networks has revolutionized the analysis and optimization of interconnected systems, enabling the development of promising and effective solutions across various technological and social domains [34]. Within the framework of complex network analysis, the identification of influential elements and the understanding of their impact on the global system have garnered significant interest in recent years. This has led to the development and evolution of various metrics and methodologies aimed at quantifying the relative importance of components within these interconnected structures.

In this context, the present review examines the development of these metrics, focusing on the progression from purely structural approaches toward more sophisticated methods that integrate both the intrinsic attributes of nodes and the weights of their connections, thus responding to the growing need for more comprehensive analyses in complex networks that better reflect real-world phenomena. The analysis encompasses both local and semi-local centrality metrics, considering the topological connections between neighbors and their relative influence within the network structure.

Consider an unweighted and undirected network $G = (V, E)$, where V represents the set of nodes and E the set of edges, depending on the application context. The adjacency matrix associated with G is described by $A = \{aij\} \in \mathbb{R}^{NxM}$, where $a_{i,j}$ represents the weight of an edge between nodes i and j. The set $T(v)$ denotes the neighbors of node v. The degree of node v, denoted as k_v, is defined according to the following Equation (4):

$$k_v = \sum_{u \in V} a_{u,v} \tag{4}$$

Each node $v \in V$ is characterized by an attribute vector $x_v = \left(x_v^1, \ldots, x_v^d\right) \in \mathbb{R}^d$, where d is the number of attributes, and each component x_v^k represents the value of the k-th attribute of node v. This characterization enables the integration of both the topological structure of the network and the intrinsic properties of its nodes in the centrality measure.

Such a description proposes a complex network characterized by weighted edges and nodes with valued attributes, allowing for the modeling of systems where the centrality and influence of each element depend on both its topological structure and the intrinsic properties of the analyzed node and those that form its relational environment. This representation is particularly relevant in contexts where the importance of an element cannot be determined solely by its connectivity patterns but requires considering the heterogeneity of nodal attributes and their interaction with the network structure.

In the context of these complex networks, where centrality depends on both the weighted topological structure and nodal attributes, the scientific literature has followed a progressive development in its approaches to measuring node importance. This development is characterized by the following three distinctive stages: initially, metrics focused exclusively on the network's topological structure, considering only node connections; subsequently, two parallel research lines emerged, one focused on incorporating edge weights and another on considering nodal attributes independently; finally, recent efforts seek to integrate both aspects into unified metrics, although this implies greater computational challenges. This evolution reflects the growing understanding of the multidimensional

nature of centrality in complex networks, where a node's importance is defined by the interaction between its structural position and intrinsic characteristics.

Table 1 presents the evolution of metrics that exclusively consider topological structure and edge weights, encompassing different network analysis levels. Among global metrics [4], Betweenness Centrality (BC) considers the frequency with which a node appears in the shortest paths between all node pairs in the network, while Closeness Centrality (CC) measures the proximity of a node to all others through geodesic distances, and Degree Centrality (DC) proposed by [35] evaluates importance according to a node's direct connections. Semi-local metrics include Local Structural Centrality (LSC) by [11], which incorporates both neighbor degrees and their local clustering coefficients, the DIL (Degree and Importance of Lines) metric by [13] that combines node degree with the weighted importance of adjacent connections, LRASP [19] which evaluates centrality considering induced subgraphs, WHC [36] that integrates multiple centrality measures, INASP [18] that combines three different aspects of local influence, K-shell Decomposition (KS) [9], and Gravity Formula Based Method (G) [37].

Table 1. Centrality metrics based on topological structure and edge weights in complex networks.

Reference	Metric	Category	Formula	Parameter Description
[35]	DC	Local	$DC(v) = k_v$	k_v is the number of neighbors of the node v.
[4]	BC	Global	$BC(v) = \sum_{s \neq v \neq t \in V} \frac{\delta_{w,v}(v)}{\delta_{w,v}}$	$\delta_{w,v}$ is the shortest path between u and w, and $\delta_{w,v}(v)$ is the shortest path between u and w passing through v.
[11]	SC	Semi-local	$SC(v) = \sum_{w \in \Gamma(v)} \sum_{w \in \Gamma(u)} k_w$	k_w is the number of neighbors of node w.
[12]	LSC	Semi-local	$LSC(v) = \sum_{w \in \Gamma(v)} \left(\alpha . k_w (1-\alpha) \sum_{w \in \Gamma(v)} C_w \right)$	α is a tunable balance parameter, and C_w is the local clustering coefficient for w.
[13]	DIL	Semi-local	$DIL(v) = k_v + \sum_{w \in \Gamma(v)} \left(I_{a_{u,v}} - \frac{k_v - 1}{k_v + k_u - 2} \right)$	$I_{a_{u,v}} = \frac{c}{\lambda}$ is the importance of $a_{u,v}$ in terms of connectivity, $\lambda = \frac{p}{2} + 1$ is the importance of $a_{u,v}$ in terms of fungibility, and p is the number of triangles with one side $a_{u,v}$.
[19]	LRASP	Semi-local	$LRASP(v) = \frac{\lvert ASP[G_{NL(v)} \backslash v] - ASPG_{NL(v)} \rvert}{ASP[G_{NL(v)}]}$	$G_{NL(v)}$ is the set of all neighbors up to level L of v in network G, and $G_{NL(v)} \backslash v$ is the induced subgraph of $G_{NL(v)}$ after node v is removed.
[36]	WHC	Semi-local	$WHC(v) = \sum_{w \in \Gamma(v)} \left(k_v + k_u + \frac{e^{ksu}/N}{EC(u)} \right) \cdot \frac{f(v,u)}{\sum_{w \in \Gamma(v)} f(v,u)}$	$EC(u)$ is the EC centrality score for node u, e denotes natural logarithm, and $f(v,u)$ is the weight between nodes v and u.
[18]	INASP	Semi-local	$INASP(v) = \alpha k_w + \beta . \sum_{l=1}^{L} \sum_{u \in V^{+l_v}} \frac{k_u}{l} + \gamma . \hat{ASP}[\hat{G_v}]$	L is the maximum neighborhood level, V^{+l_v} is the set of nodes at $l - hop$ from v, and $\hat{ASP}[\hat{G_v}]$ is defned as $LRASP$. Additionally, α, β, and γ are the impact coefficients.
[9]	KS	Global	$Ks(v_i) = \max\{k \mid v \in S_k\}$	k is the k − core of the network, the set of remaining nodes after iteratively removing all nodes with degree $\leq k$.
[37]	G	Semi-local	$G(v_i) = \sum_{uj \in \Psi(v_i)} \frac{k_s(v_i) . k_s(v_j)}{S^2(v_i, v_j)}$	L is the considered neighborhood level (typically $l = 3$). The set $\Psi(v_i)$ contains nodes up to L levels of distance.

More advanced metrics focus on information propagation and node distance in complex networks, such as LARSP, ARP, and LASP. LARSP (Local Average Shortest Path) is a local metric that measures node centrality based on the average length of shortest paths

from one node to all other nodes in its local subgraph. Its objective is to capture the node's influence on information propagation within its immediate neighborhood, considering how local connectivity impacts the node's capacity to transmit information across the network. ARP (Average Reciprocal Path) extends the LARSP concept by considering the reciprocal distance of shortest paths in a directed network. Specifically, it evaluates how the path structure between nodes, considering edge directions, affects node centrality. LASP (Local Average Shortest Path) is an optimized version of LARSP that incorporates a weighted local average of shortest distances, reducing computational complexity by focusing on each node's local subgraph.

Meanwhile, metrics considering valued nodal attributes (Table 2) also present different analytical scopes. Local metrics include node-weighted degree [7], which modifies the traditional degree definition by incorporating a nodal weight function, and WNDegree variants [8] that integrate nodal attributes with local topological structure. Semi-local metrics include node-weighted harmonic centrality [14] that considers geodesic distances in the extended neighborhood, node-weighted betweenness centrality [15] that incorporates the importance of communication between nearby node pairs, and modified eigenvector centrality [16] that adjusts nodal weight influence through a variable parameter. Additionally, hybrid metrics have been developed, such as the nodal attribute screening method, applicable at both local and global levels, and the MCNDI metric that integrates multiple indicators through the CRITIC method, combining local and global aspects [17]. The Node Attribute Screening Centrality method [38] uses a regression model in which the centrality is the dependent variable, the nodal attributes are the predictor variables, and the regression coefficients represent the influence of each attribute.

The weighted K-shell method [39] uses the node degree and link weights, while the weighted K-shell degree neighborhood method [40] combines the degree and the k-shell index with adjustable parameters.

The integration of nodal attributes and connectivity in combined centrality metrics began with seminal works, such as [41], who addressed attributed graph analysis by incorporating categorical attributes in centrality evaluation. Their proposal extends classical measures through the E-I homophily index and betweenness metrics, enabling node classification into groups based on qualitative characteristics. While this approach represents an initial step in considering nodal attributes, it is limited to categorical characteristics without exploiting the richness of numerical attributes that could more precisely capture actors' influence in the network.

A more comprehensive advancement in integrating global structure and attributes was proposed by [8], who developed a metric called node and edge-weighted closeness centrality, which calculates nodal importance considering both normalized distances between nodes and connection weights along shortest paths. This measure integrates the network's global structure and connection weights into a global centrality metric. This measure, denoted as $CWNECentr(v_i) = w_i \cdot CWECentr(v_i)$, is defined as the product between the weight of node v_i and its weighted closeness centrality $CWECentr(v_i)$. The latter is calculated as $CWNECentr(v_i) = \sum_{u_j \in V\{i\}} \frac{\sum_{e \in ShorthPath(v_i,\, v_j)} w(e)}{|ShortParth(v_i,\, v_j)|}$, where $w(e)$ represents the edge weight e in the shortest path, and $ShortParth(v_i,\ v_j)$ denotes the length of said path, measured as the number of links between nodes v_i and v_j. However, its main limitation lies in the need to calculate the shortest distances between each node pair, resulting in high computational complexity, especially in extensive and complex networks. This complexity increases significantly in networks with weights and nodal attributes, due to the additional analysis required for each connection.

The development of centrality metrics reflects a progression from purely structural approaches toward approximations that incorporate edge weights or nodal attributes inde-

pendently. However, there exists a significant gap in developing metrics that simultaneously integrate both edge weights and valued nodal attributes while maintaining manageable computational complexity. Existing attempts, such as that in [8], although promising, face significant limitations in terms of scalability and computational efficiency. This gap is particularly relevant in the current context, where complex networks frequently exhibit heterogeneity in their connections and diversity in their node characteristics. Therefore, developing a centrality metric that can efficiently capture this duality while maintaining feasible computational complexity represents a necessary research direction to advance the understanding and analysis of real-world complex networks.

Table 2. Node-attribute-based centrality metrics.

Reference	Metric	Category	Formula	Parameter Description
[7]	Node-weighted degree centrality	Local	$\sum_{u \in V(u)} \frac{f(w_x).a_{u,v}}{\sum f(w_x)}$	$f(w_x)$ is the weight function of node x.
[14]	Node-weighted harmonic centrality	Semi-local	$f(w_x) + \frac{\sum_{u \in V(u)} \frac{f(w_u)}{d_{u,v}+1}}{\sum_{x \in V} f(w_x)}$	$d_{u,v}$ is the shortest path distance between nodes u and v.
[15]	Node-weighted betweenness centrality	Semi-local	$\sum_{s,t \in ; s \neq u \neq t} f(w_s, w_t) \frac{\sigma_{st}(u)}{\sigma_{st}}$	$\sigma_{st}(u)$ is the number of shortest paths between s and t that pass through u.
[16]	Modified node-weighted eigenvector centrality	Semi-local	$EC(u) \cdot W_u^{\beta}$	$\beta \in [-1, 1]$ is an adjustment parameter, and $EC(u)$ is the unweighted eigenvector centrality.
[8]	Node-weighted degree centrality	Local	$WNC(v_i) = w_i \left[Degree(v_i) + \alpha . \left((\text{degout}(v_i))^{1-\alpha} . \sum_{vjeout(v_i)} w_{i,j}^{\alpha} \right) \right]$	w_i represents the weight of node v_i, while $Degree(v_i)$ denotes the degree of the node. $degout(v_i)$ refers to the number of outgoing connections of the node. The parameter α controls the balance in the equation, and $w_{i,j}$ represents the weight of the link between nodes v_i and v_j.
[17]	Multi-attribute CRITIC network decision indicator (MCNDI)	Multi-attributes	$MCNDI_i = w_1 \cdot H_i + w_2 \cdot COC_i + w_3 \cdot KS_i + w_4 \cdot NCC_i$	H_i represents the $H-index$, COC_i denotes closeness centrality, KS_i corresponds to the $k-shell$ value, and NCC_i signifies the network constraint coefficient. The weights w_j are computed using the CRITIC method.
[38]	Node attribute screening centrality	Local/Global	$y = b_0 + b_1 x_1 + \ldots + b_i x_i$	y is a centrality, x_i is a nodal attribute, and b_i is a regression coefficient.
[39]	Weighted k-shell	Local	$k_w^e = \lambda k_v + (1+\lambda) \sum_{uj \in \Gamma_i} w(v_i, v_j)$	λ is a parameter between 0 and 1 that balances the importance between the node degree and the potential weights of the links.
[40]	Weighted k-shell degree neighborhood	Semi-local	$\text{ksd}^n(v_i) = \sum_{uj \in \Gamma(v_i)} w_{ij}$ $w_{ij} = (c_1.k(v_i) + c_2.k_s(v_i) + c_3.k(v_i) + c_4.k_s(v_i))$	c_1, c_2, c_3, c_4 are four adjustable parameters between 0 and 1 that control the relative contribution of the degree and the k-shell index of the source and target nodes.

3.2. Proposed Metric: Semi-Local Centrality with Weighted and Lexicographic Extended Neighborhood in Node-Attributed Weighted Networks (SL-WLEN)

The SL-WLEN metric quantifies node centrality in complex networks based on the LARSP (Local Average Shortest Path) connectivity analysis through its DegreeLocal and

DegreeSemiLocal components, which evaluate partial centrality as a function of connection degrees. SL-WLEN extends this foundation by incorporating the following two additional components: a local component through the normalized node value, and a semi-local component via SemilocalNodeLexOrder, which introduces the lexicographic ordering of neighbors. This component combination enhances the metric's capability to reflect the influence of characteristics in chip manufacturing, enabling the identification of the most relevant features of the final product by considering both their connectivity and their intrinsic values, as well as their structural position within the network. Figure 1 illustrates the metric implementation process.

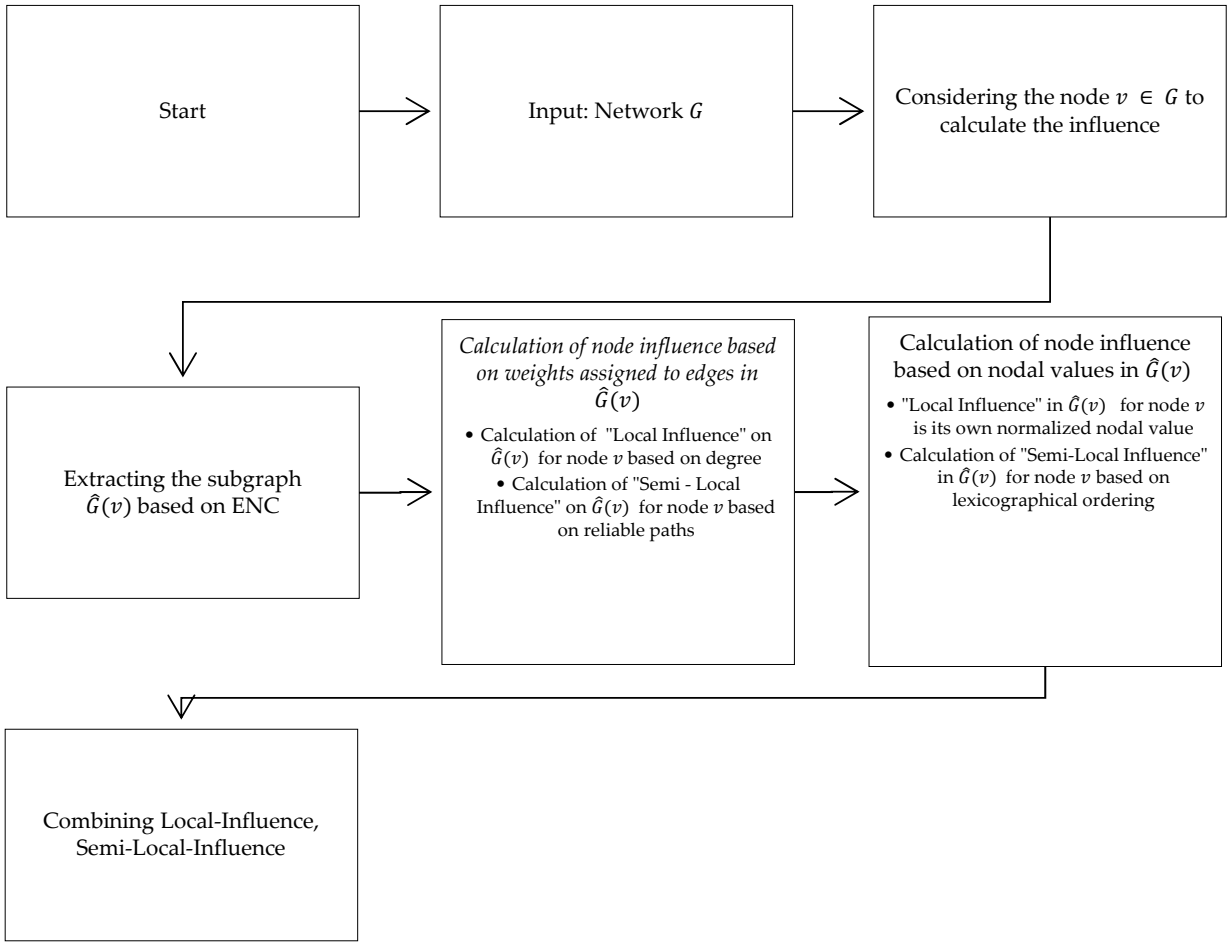


Figure 1. General implementation process of the proposed centrality metric.

3.3. Integration of Lexicographic Ordering

SL-WLEN integrates SemilocalNodeLexOrder, enabling a more precise characterization of node influence within its structural and attributive context. The implementation of SL-WLEN is based on Extended Neighborhood Connectivity (ENC), which extracts a subgraph encompassing node neighbors up to distance L. For details on ENC, see the work of [21], pages 114 and 115.

Once the subgraph is obtained through ENC, SemilocalNodeLexOrder quantifies node influence by considering its position in a lexicographic ordering based on attributes and neighborhood structure. At each distance level l, SemilocalNodeLexOrder assigns higher weights to better-positioned nodes within the ordering, allowing the capture of subtle differences in nodes' relative importance.

The metric operates by considering (1) the prioritization of important features through lexicographic comparison, (2) influence penalization as distance increases and the adjustment of node influence based on neighbor connectivity, and (3) influence accumulation.

The prioritization of important features through lexicographic comparison is discussed as follows:

In the chip quality network, each node represents a quality characteristic, and its importance depends not only on its individual contribution but on its relationship with other characteristics. The SemilocalNodeLexOrder function enables node ordering based on their relative importance within their neighborhood, ensuring that the most influential characteristics maintain a priority position. Given node v, its local influence is measured from the lexicographic ordering of its immediate neighborhood at distance l. The set of neighbors $N_l(v)$ is ordered according to the importance value of each characteristic $val(u)$, obtaining $OrderLex_l(v) = [u_1, u_2, \ldots, u_k]$ such that $\forall i < j,\ val\ (u_i) \geq al\ \left(u_j\right)$. This ordering favors nodes with highly relevant characteristics for chip quality, ensuring that those with higher values carry greater weight in the metric. In terms of chip manufacturing, this means that characteristics that most influence defects or improvements in the final product will occupy priority positions within the centrality evaluation. The partial contribution of node v at level l is defined as $\phi_l(v) = \frac{(N-pos(OrderLex_l(v)))}{N-1}$, where $N - pos(OrderLex_l(v)$ is v's position in the lexicographic order. If a node has neighbors with a high impact on chip quality, its position in the list will be lower (closer to 1), increasing the numerator and, consequently, its influence in the metric.

Distance-based influence penalization and node influence adjustment based on neighbor connectivity are discussed as follows:

In the chip quality network, the effect of a characteristic can propagate through multiple interactions. However, its impact must be reduced with distance to prevent the overvaluation of distant connections. The influence of v at each level l is weighted according to its neighborhood size and the maximum connectivity at that level $\phi_l(v) = \frac{(N-pos(OrderLex_l(v)))}{N-1} \cdot \frac{|N_l(v)|}{\max\{|N_l(x)|:x\in V\}} \cdot \frac{1}{L+l}$. Here, the first term maintains the lexicographic priority based on the characteristic's importance, while the second term adjusts the relative contribution according to neighborhood size, enabling the differentiation of highly connected characteristics, and the third term introduces a penalization that reduces influence as distance increases, modeling the decreasing effect of characteristic propagation in manufacturing.

This adjustment aims to capture indirect relationships between characteristics without excessively diluting or overestimating their influence, ensuring that closer nodes have a more relevant impact on the metric, while the effects of distant nodes are attenuated in a controlled manner. In the context of chip quality, this approach helps evaluate not only directly influential characteristics but also those affecting the product in a more indirect yet equally relevant way, without excessive overvaluation.

Influence accumulation is discussed as follows:

Finally, the total semi-local influence of node v is obtained by accumulating partial contributions at each exploration level up to maximum L in the form $\varphi SemilocalNodeLexOrder(v) = \sum_{l=1}^{L} \phi_l(v)$. This enables the consideration of how a characteristic affects chip quality not only directly but also through indirect relationships with other characteristics. Additionally, it balances influence from $l = 1$ to $l = L$, preventing nodes with high connectivity from dominating the metric and providing a fair evaluation based on network structure. Lexicographic ordering proves particularly appropriate for evaluating chip manufacturing quality due to its unique capability to preserve the importance of critical characteristics. Unlike existing metrics, which tend to dilute the influence of important characteristics through

various procedures, lexicographic ordering maintains the relevance of the most significant nodes throughout the analysis.

Traditional metrics present limitations in this context. Some use weighted sums like node-weighted degree centrality, others rely on distance normalizations like node-weighted harmonic centrality, or they employ shortest paths like node-weighted betweenness centrality. There are also those that apply products with adjustable parameters, linear regressions, or a combination of multiple indices, such as MCNDI. All these approaches may inadvertently reduce the influence of critical characteristics through their statistical aggregations. In contrast, lexicographic ordering preserves the relative importance of each characteristic through three complementary aspects; it prioritizes nodes based on their individual value, connectivity level, and influence adjusted by distance. This combination enables a more precise evaluation, where the importance of each characteristic is determined by its own value and its relationships with neighboring characteristics, without losing critical information in the process.

3.4. Definitions

The SL-WLEN metric quantifies node centrality in a complex network by considering two levels of analysis, local and semi-local, and integrating weighted connectivity components and nodal attributes. Its purpose is to capture node influence not only through direct connectivity but also by evaluating the importance of its neighbors at different proximity levels, their characteristics, and their relative position in the network. To achieve this, it integrates four main factors, local influence by connectivity (DegreeLocal), local influence by node (ζV_v^{norm}), semi-local degree influence (DegreeSemiLocal), and semi-local node value influence based on lexicographic ordering (SemilocalNodeLexOrder).

In the final metric, (ζV_v^{norm}) normalizes the node value by relativizing it within its neighborhood, capturing its intrinsic importance beyond structural connectivity.

Definition 1. *Local influence by connectivity (DegreeLocal).*

DegreeLocal captures the local influence of the node based on its direct connectivity, normalizing the node degree with respect to the total network size. This reflects its immediate importance within the network. The local influence by the connectivity of v denoted as $\phi DegreeLocal(v)$ is defined according to the following Equation (5):

$$\varphi Degree\ Local(v) = \frac{k_v}{N} \tag{5}$$

where k_v is the degree of node v, and N is the number of directly connected nodes.

Definition 2. *Semi-Local Degree Influence (DegreeSemiLocal).*

DegreeSemiLocal, derived from LARSP [20] and based on LASP [21], quantifies semi-local influence by considering nearby neighbors within a subgraph extracted through the ENC (extended neighborhood connectivity) concept. This influence is weighted based on several aspects, including weighted connectivity, which reflects the intensity of relationships between the node and its neighbors through edge weights, proximity, where neighbor influence decreases as distance increases, modeling impact propagation within the network, and structural importance, which prioritizes neighbors with higher topological relevance.

The semi-local influence of v, denoted as $\varphi DegreeSemiLocal(v)$, is defined according to the following Equation (6):

$$\varphi Degree\ Semi\ Local(v) = \frac{1}{|G_{NL}(v)|}\sum_{l=1}^{L}\sum_{u\in T_l(v)}\sqrt{\frac{w_{u,v}.k_v}{d_{u,v}(k_u + k_{v)}}} \tag{6}$$

where $|G_{NL}(v)|$ is the set of all neighbors up to level L of node v in network G, and $T_t(v)$ is the set of all neighbors at level l of node v.

Definition 3. *Semi-Local Node Value Influence Based on Lexicographic Ordering (SemilocalNodeLexOrder).*

SemilocalNodeLexOrder introduces a novel perspective through the lexicographic ordering of nodes based on their attributes and neighborhood structure. This evaluates how a node's relative position within this order affects its influence, considering its structural and attributive environment at different distance levels. Additionally, it includes the contribution of nearby neighbors within a subgraph extracted through ENC, enabling a deeper evaluation of the node within its topological and attributive context.

The semi-local node influence based on lexicographic *Ordering* $\varphi SemilocalNodeLexOrder(v)$ measures node influence by considering its nearby neighbors at different distance levels (up to a maximum L). Each level contributes with a partial contribution ϕ_l that depends on the node's position in the lexicographic order $OrderLex_l(v)$ within its neighborhood, the number of neighbors at that distance, and the maximum degree among nodes at the same level according to the following Equation (7):

$$\varphi SemilocalNodeLexOrder(v) = \sum_{l=1}^{L}\phi_l(v) \tag{7}$$

where ϕ_l the partial contribution per level is defined according to the following Equation (8):

$$\phi_l(v) = \frac{(N - pos(OrderLex_l(v)))}{N-1}.\frac{|N_l(v)|}{max\{|N_l(x)|: x \in V\}}.\frac{1}{L+l} \tag{8}$$

where $N = |V|$ is the total number of nodes, $pos(OrderLex_l(v))$ defines node v's position in the lexicographic order $OrderLex_l$ at level l, $|N_l(v)|$ is the number of neighbors at distance l from node v, $\max\{|N_l(x)|: x \in V\}$ is the maximum degree among all nodes at level l, and L is the maximum exploration level.

The ordering function $OrderLex_l(v)$ is defined as an ordered set of nodes based on a lexicographic comparison, as shown in the following Equation (9):

$$OrderLex_l(v) = [u_1, u_2, \ldots, u_k | u_i \in V,\ \forall i, j:\ Comp_l(u_i, u_j) = 1 \tag{9}$$

While $\phi Comp_l(u, z)$:

$$\phi Comp_l(u,z) = \begin{cases} 1,\ si\ \exists\ i \le min(|N_l(u)|, |N_l(z)|) : \left(\forall j < i : val(N_l(u))_j = val(N_l(z))_j \wedge val(N_l(u))_j > val(N_l(z))_i\right) \\ 0,\ otherwise \end{cases} \tag{10}$$

where $N_l(v) = \{u \in V | d(v,u) = 1\}$ is a set of neighbors of v at distance l, $\text{val}(N_l(v)) = \text{sortdesc}(\{\text{val(u)}|\text{u} \in N_l(u)\})$ are values of v's neighbors sorted in descending order, and $d(v,u)$ is the shortest path length between v and u in G.

This metric uses normalized values for each node, obtained by dividing its value by the maximum value of its neighbors at the same distance level, which adjusts its influence based on relative importance within the neighborhood.

3.5. Special Considerations

If several nodes have the same lexicographic order ($Comp_l(u,z) = Comp_l(z,\ u) = 0$, then it is resolved by assigning the same order for the set of nodes V. The term $\varphi Semilocal NodeLexOrder(v)$ can be interpreted as the sum of contributions $\phi_l(v)$ from each level l, facilitating a detailed analysis of each node's behavior at each exploration level.

Definition 4. *Total Influence.*

For a node v, SL-WLEN is defined by the following Equation (11):

$$SL - WLEN(v) = \xi.\varphi DegreeLocal(v) + \zeta V_v^{norm} + \delta.\varphi Degree\ Semilocal(v) + \gamma \varphi SemilocalNodeLexOrder(v) \quad (11)$$

where ξ, ζ, δ, and γ are adjustable parameters between 0 and 1. The first two control the local and semi-local influence of node connectivity, and the remaining ones control the local and semi-local influence of its nodal value, satisfying the condition $\xi + \zeta + \delta + \gamma = 1$.

4. SL-WLEN Example

To better clarify the computational procedure of the proposed metric, we describe a numerical example. An undirected weighted graph with 11 nodes and 14 edges is assumed, as shown in Figure 2. We present a calculation example for v_6, considering $L = 2$ and edge weights $w_{u,v}$ as shown on the edges, with assigned nodal values $V = \{V_{1,2,9,11} = 1.5; V_{3,4,5,7} = 0.5;\ V_{6,8,10} = 4\}$.

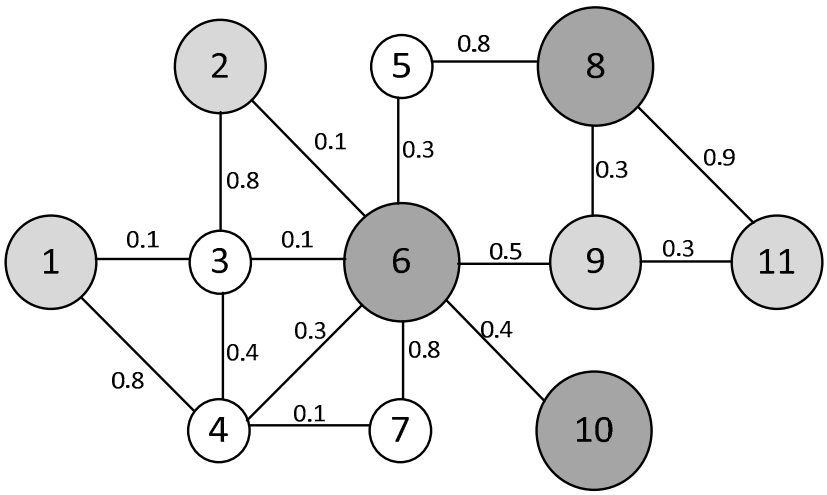


Figure 2. A simple graph with 11 nodes and 14 edges.

According to Definition 1, $k_6 = 7$ and $N = 11$. According to Definition 2, the calculation of $\varphi Degree\ Semi\ Local(6)$ is performed as follows:

$$\begin{aligned}\varphi Degree\ Semi\quad & Local(6)\\ &= \tfrac{1}{10}\left[\left(\sqrt{\tfrac{0.1\ (7)}{2+7}}\right)_{u=2}+\left(\sqrt{\tfrac{0.1(7)}{4+7}}\right)_{u=3}+\left(\sqrt{\tfrac{0.3\ .(7)}{4+7}}\right)_{u=4}+\left(\sqrt{\tfrac{0.3(7)}{2+7}}\right)_{u=5}\right.\\ &\left.+\left(\sqrt{\tfrac{0.8\ (7)}{2+7}}\right)_{u=7}+\left(\sqrt{\tfrac{0.5(7)}{3+7}}\right)_{u=9}+\left(\sqrt{\tfrac{0.4\ (7)}{1+7}}\right)_{u=10}\right]_{l=1}\\ &+\left[\left(\sqrt{\tfrac{0.3\ (7)}{2(2+7)}}\right)_{u=1}+\left(\sqrt{\tfrac{0.3(7)}{2(3+7)}}\right)_{u=8}+\left(\sqrt{\tfrac{0.3\ .(7)}{2(2+7)}}\right)_{u=11}\right]_{l=2}=0.4028\end{aligned} \tag{12}$$

$\varphi SemilocalNodeLexOrder(6)$ is determined from Definition 3.

At level 1, the neighbors of v_6 are nodes number 2, 3, 4, 5, 7, and 9, giving a total of $||\ N_1(6)|=7|$ neighbors. To calculate the lexicographic order, the normalized values associated with these nodes are considered, which are 0.375, 0.125, 0.125, 0.125, 0.125, 0.375, and 1.00. These values are sorted in descending order to form node V6's signature, resulting in {1.0, 0.375, 0.375, 0.125, 0.125, 0.125, 0.125}. Comparing this signature with other nodes in the graph generates a ranking where nodes with higher signatures are placed first. In this case, node v_6 occupies position 4 in the level 1 lexicographic ranking, defined as $OrderLex_1$.

At level 2, the neighbors of v_6 are nodes number 1, 8 and 11, resulting in $|\ N_2(6)|=3|$ neighbors. The normalized values associated with these neighbors are 0.375, 1.0, and 0.375, which are sorted in descending order to form v_6's signature at this level, obtaining $\{1.0,\ 0.375\ y\ 0.375\}$. Following the same signature comparison process to determine lexicographic order, it is concluded that v_6 occupies position 4 in the level 2 lexicographic ranking, defined as $OrderLex_2$. Substituting the values in Equations (6) and (7), $\varphi SemilocalNodeLexOrder(6)$ is determined as follows:

$$\varphi SemilocalNodeLexOrder(6)=\left[\frac{(N-pos_1)}{(N-1)}\cdot\frac{|N_1(6)|}{max\{|N_1|\}}\cdot\frac{1}{L+l}\right]_{l=1}+\left[\frac{(N-pos_2)}{(N-1)}\cdot\frac{|N_2(6)|}{max\{|N_2|\}}\cdot\frac{1}{L+l}\right]_{l=2} \tag{13}$$

$$\varphi SemilocalNodeLexOrder(6)=\left[\frac{(11-4)}{(11-1)}\cdot\frac{7}{7}\cdot\frac{1}{2+l}\right]_{l=1}+\left[\frac{(11-4)}{(11-1)}\cdot\frac{3}{7}\cdot\frac{1}{2+2}\right]_{l=2} \tag{14}$$

$$\varphi SemilocalNodeLexOrder(6)=0.2333+0.0750=0.3083 \tag{15}$$

Finally, SL-WLEN(3) is calculated according to the adjustable parameters condition $\xi=\zeta=\delta=\gamma=0.25$, in accordance with the following Equation (16):

$$SL-WLEN\ (6)=(0.25\times0.636)+(0.25\times1)+(0.25\times0.403)+(0.25\times0.30)=0.587 \tag{16}$$

The SL-WLEN metric results for all nodes are shown in Table 3.

Table 3. SL-WLEN metric results for all nodes in the example network.

Rank	Node	φDegreeLocal	ζV_v^{norm}	Degree Semilocal	SemilocalNodeLexOrder(v)	SL-WLEN(v)
1	V6	0.636	1.000	0.403	0.308	0.587
2	V8	0.273	1.000	0.376	0.025	0.419
3	V10	0.091	1.000	0.166	0.010	0.317
4	V9	0.273	0.375	0.285	0.171	0.276
5	V4	0.364	0.125	0.343	0.229	0.265
6	V2	0.182	0.375	0.191	0.231	0.245
7	V3	0.364	0.125	0.262	0.210	0.240
8	V5	0.182	0.125	0.262	0.345	0.228
9	V11	0.182	0.375	0.297	0.055	0.227
10	V1	0.182	0.375	0.200	0.054	0.203
11	V7	0.182	0.125	0.223	0.157	0.172

5. Experimental Results

In Figure 3, a general view of the complex network for chip manufacturing quality control is presented. The visualization shows the complete network structure, where nodes (circles) represent quality characteristics, and edges (lines) represent the correlations between them.

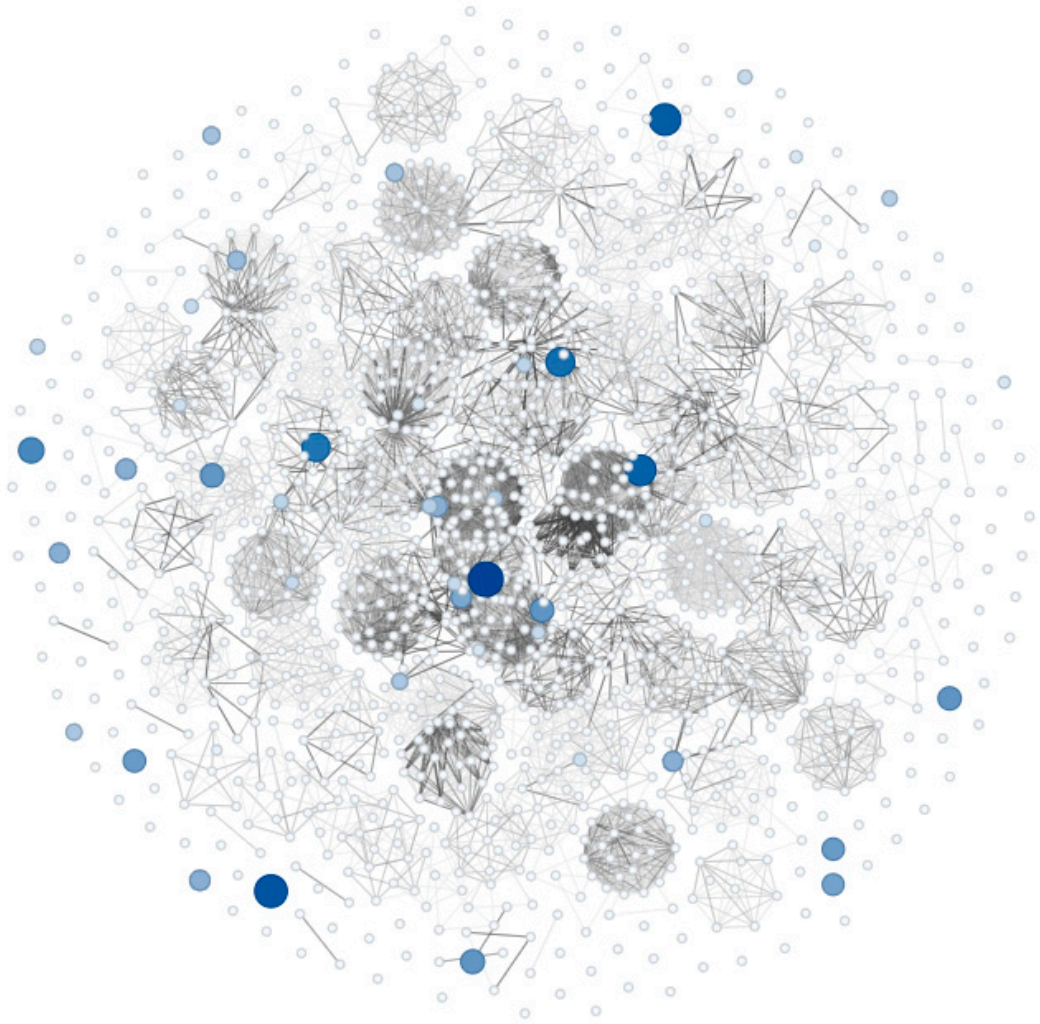


Figure 3. Overview of the complex network for chip manufacturing quality control, where node size and blue intensity represent the characteristic importance values derived from the logistic regression model with L1 (Lasso) regularization.

Figure 4 presents the detailed visualization of the quality control network. Node size and blue color intensity indicate the individual importance level of each characteristic (nodal value); larger size and darker blue tonality correspond to greater importance. Connections between nodes (edges) are represented on a grayscale, where tonalities closer to black indicate stronger correlations between characteristics, while lighter tones represent weaker correlations.

Figure 5 presents the visualization of the quality control network with characteristic identifiers. The features are identified with the prefix "f" followed by a four-digit sequential number. For example, f1263 corresponds to feature number 1263 of the process.

Figure 6 illustrates the visualization of the network's structural configuration in core and peripheral zones. Panel (a) shows how the network accumulates connections in high-density areas, with strongly interconnected nodes forming clusters that reveal grouping patterns from the network's center outward. Meanwhile, panel (b) illustrates the net-

work's peripheral region, where nodes with lower connectivity are located, demonstrating how these elements are spatially distributed in areas furthest from the network's center. This progressive representation facilitates an understanding of the network's complexity from different perspectives, enabling the direct appreciation of the relationships between process characteristics.

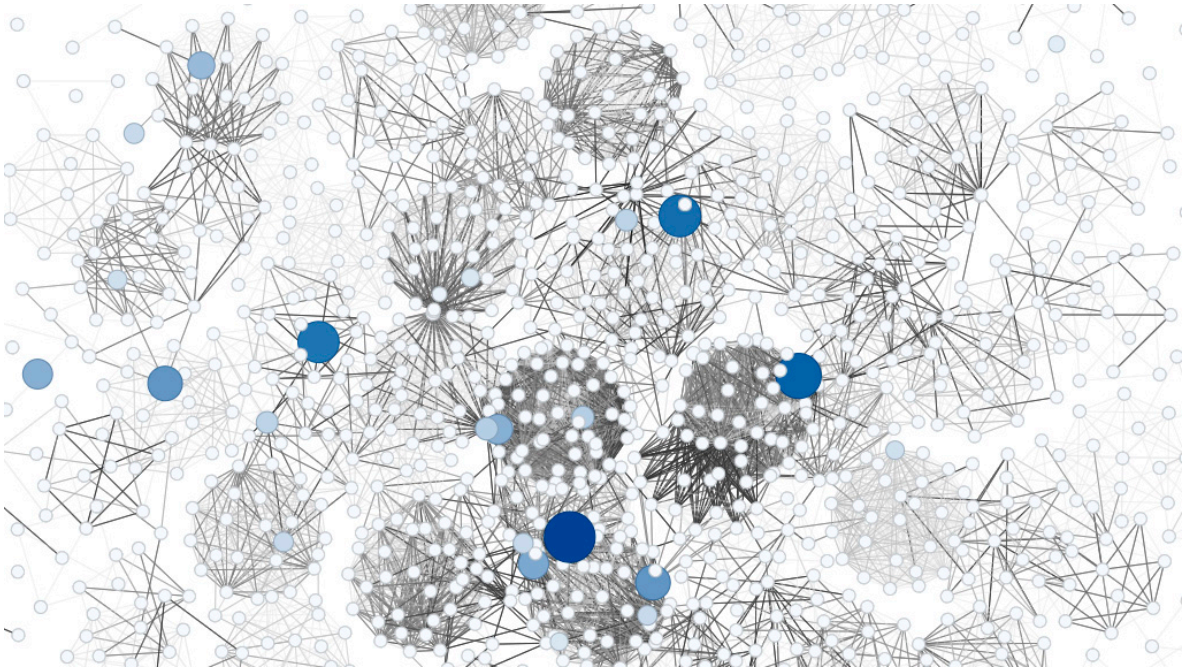


Figure 4. Detailed representation of the quality control network in chip manufacturing, where node diameter and blue intensity reflect L1-regularized importance values.

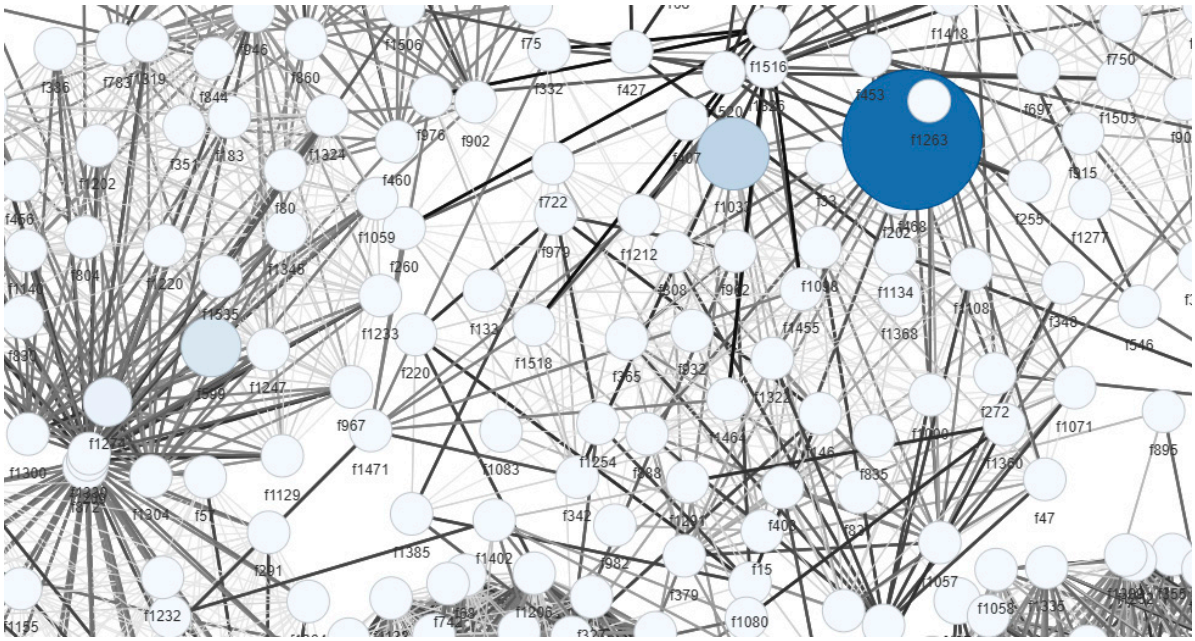
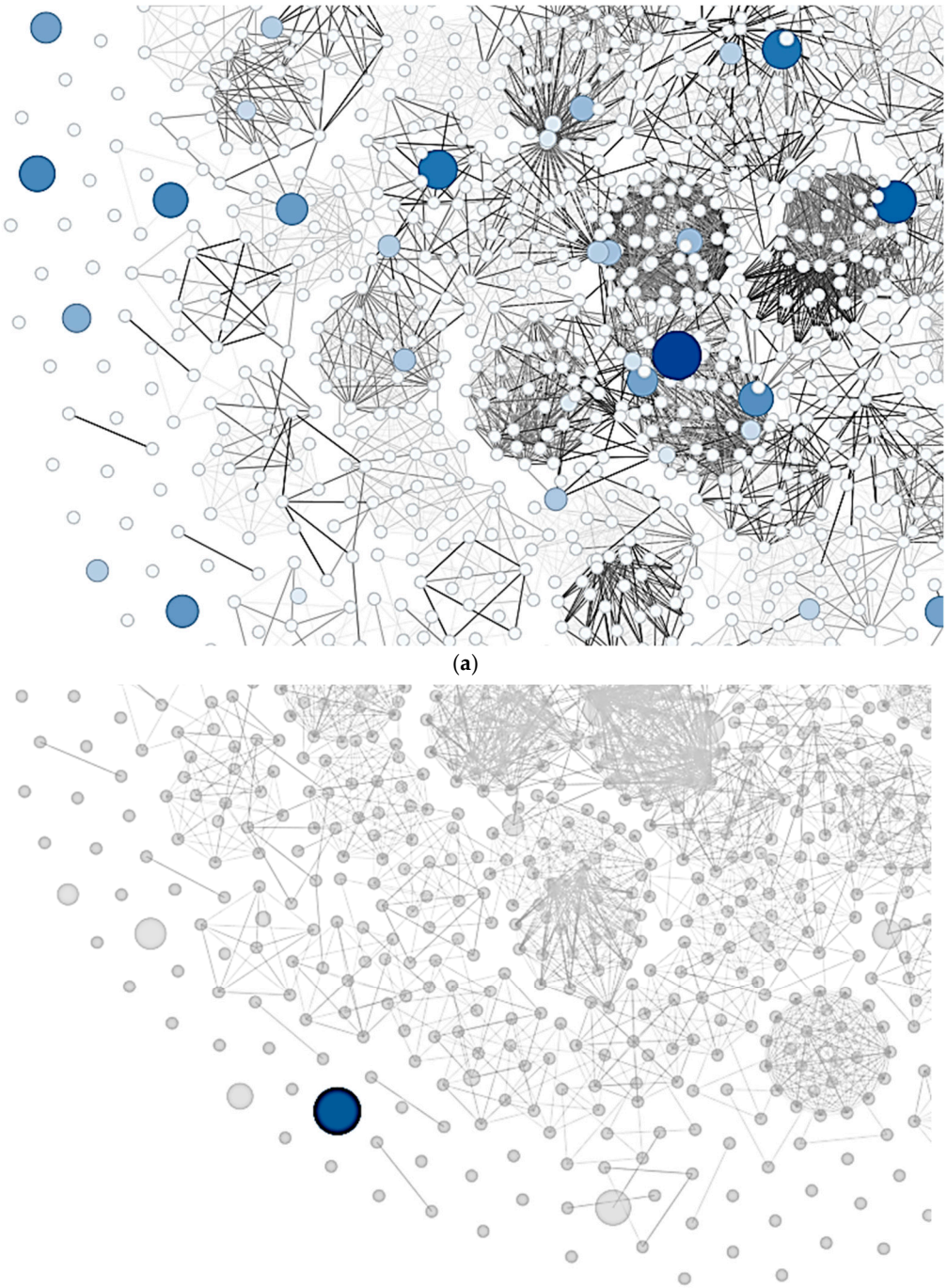


Figure 5. Visualization of the quality control network with feature identifiers. Node diameter and blue intensity reflect L1-regularized importance values.

The visualization of the quality control network in chip manufacturing maintains graphical legibility, enabling the identification of characteristics' importance hierarchy through node size and tonality, as well as correlation strength through connection intensity.

The representation achieves a balance between showing densely connected structures (clusters) and more dispersed zones. This clarity in visualization facilitates an understanding of the complex network of interrelationships in the manufacturing process, providing an effective visual tool for quality control monitoring and analysis.



(**a**)

(**b**)

Figure 6. Visualization of network structural configuration, core, and peripheral zones. Panel (**a**) shows central nodes with high connectivity while panel (**b**) displays peripheral nodes with low connectivity.

Table 4 shows the top 20 nodes with highest centrality according to the SL-WLEN metric, including their components and final rankings.

Table 4. SL- Top 20 SL-WLEN metric applications in measuring quality characteristic centrality.

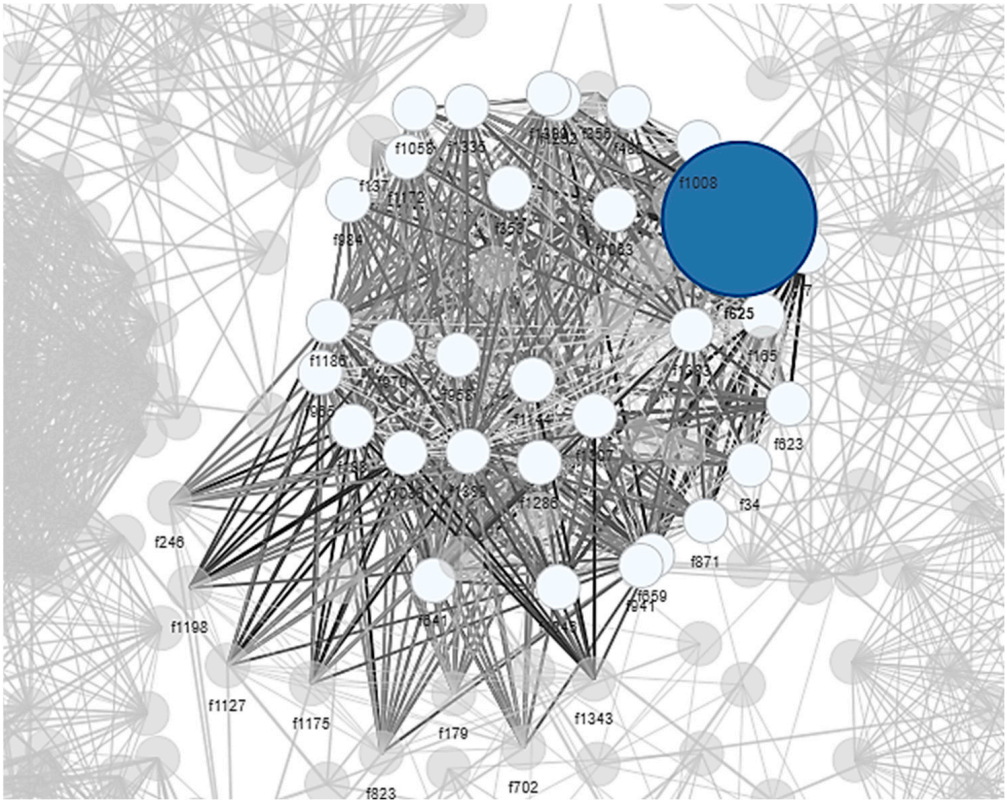
Node	φDegreeLocal	ζV_v^{norm}	Degree Semilocal	SemilocalNodeLexOrder(v)	SL-WLEN(v)	Rank
f625	0.0193	0.8412	0.6195	0.3047	0.4462	1
f1397	0.0148	1.0000	0.6081	0.1095	0.4331	2
f468	0.0096	0.7487	0.3923	0.3509	0.3754	3
f506	0.0019	0.5576	0.4357	0.3796	0.3437	4
f981	0.0109	0.4842	0.6195	0.2365	0.3378	5
f732	0.0277	0.3096	0.6149	0.3245	0.3192	6
f1020	0.0167	0.3912	0.6026	0.2623	0.3182	7
f18	0.0051	0.5979	0.5259	0.1407	0.3174	8
f901	0.0148	0.4201	0.4320	0.3188	0.2964	9
f1245	0.0193	0.2558	0.6593	0.2277	0.2905	10
f1048	0.0039	0.2307	0.5820	0.3438	0.2901	11
f396	0.0006	0.7429	0.3657	0.0417	0.2877	12
f181	0.0013	0.9256	0.0000	0.1667	0.2734	13
f1165	0.0026	0.4705	0.5363	0.0833	0.2732	14
f7	0.0193	0.0000	0.6410	0.4116	0.2680	15
f1209	0.0251	0.0000	0.7026	0.3346	0.2656	16
f1033	0.0103	0.2543	0.5077	0.2681	0.2601	17
f721	0.0109	0.0000	0.5159	0.4992	0.2565	18
f84	0.0161	0.0776	0.5680	0.3598	0.2553	19
f1176	0.0257	0.0000	0.5118	0.4815	0.2548	20

Figure 7 presents visual local subnetworks corresponding to the six highest-ranked nodes according to the SL-WLEN metric, revealing distinctive patterns of connectivity and local structure. The composite visualization shows different topological configurations that justify the ranking obtained through the proposed metric.

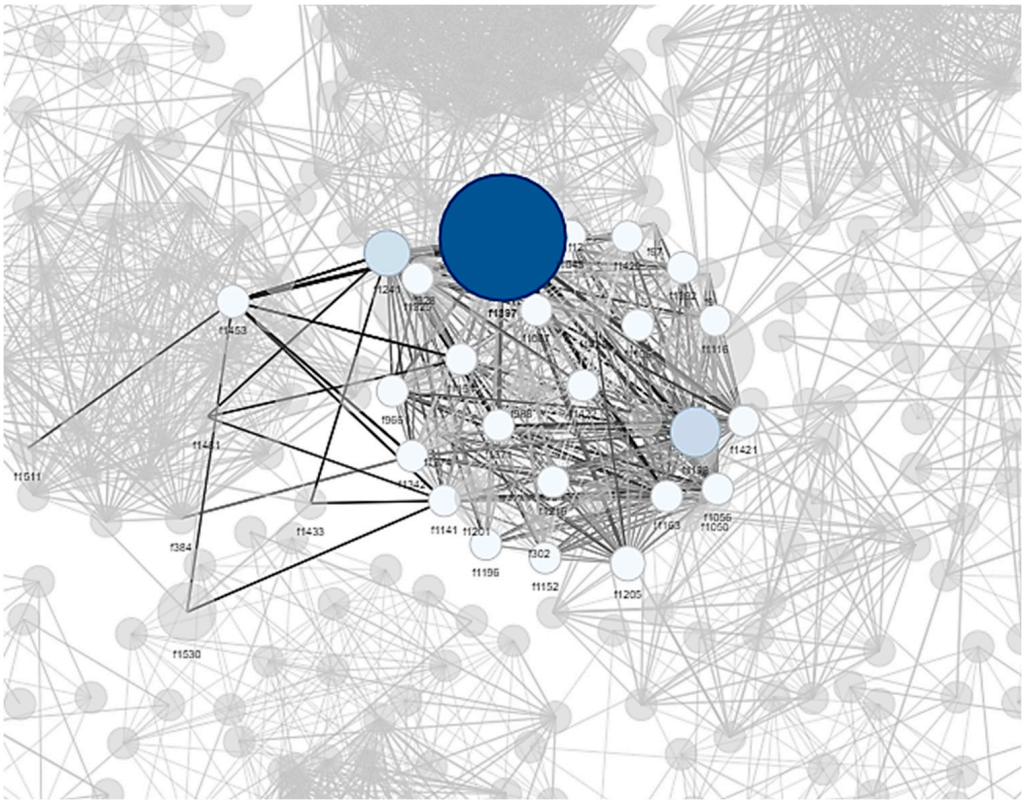
Node f625, which occupies the first position, exhibits the high density of local connections with a compact and well-connected structure, characterized by multiple intermediate nodes forming a cohesive community. The second highest-ranked node, f1397, presents a distinctive triangular connectivity pattern, less dense than f625 but with strategically distributed connections in its neighborhood. In the third position, f468 shows a predominantly radial structure with direct connections and a more pronounced dispersion pattern than the previous ones. Node f506, in fourth place, is characterized by minimal but strategic connectivity, with sparse links and a simpler structure compared to higher-ranked nodes. The fifth node, f981, presents moderate connection density with a semi-compact structure and irregular link distribution. Finally, f732, in sixth position, is distinguished by a hexagonal structure with regular and symmetric connections, showing moderate density with an ordered pattern.

This visualization provides empirical evidence of how the SL-WLEN metric captures different aspects of centrality and local structure in the network. Visually, it is possible to appreciate its capability to identify significant nodes based on multiple topological and structural criteria.

The visualizations are consistent with the SL-WLEN ranking. The following top three nodes (f625, f1397, f468) demonstrate more sophisticated connectivity patterns that reflect their high metric values: f625 with its dense and cohesive structure ($SL - WLEN = 0.4462$), f1397 with strategic triangular connections ($SL - WLEN = 0.4331$), and f468 with its efficient radial pattern ($SL - WLEN = 0.3754$). The lower-ranked nodes (f506, f981, f732) exhibit simpler or less-integrated structures, consistent with their lower SL-WLEN values (0.3437, 0.3378, and 0.3192, respectively).

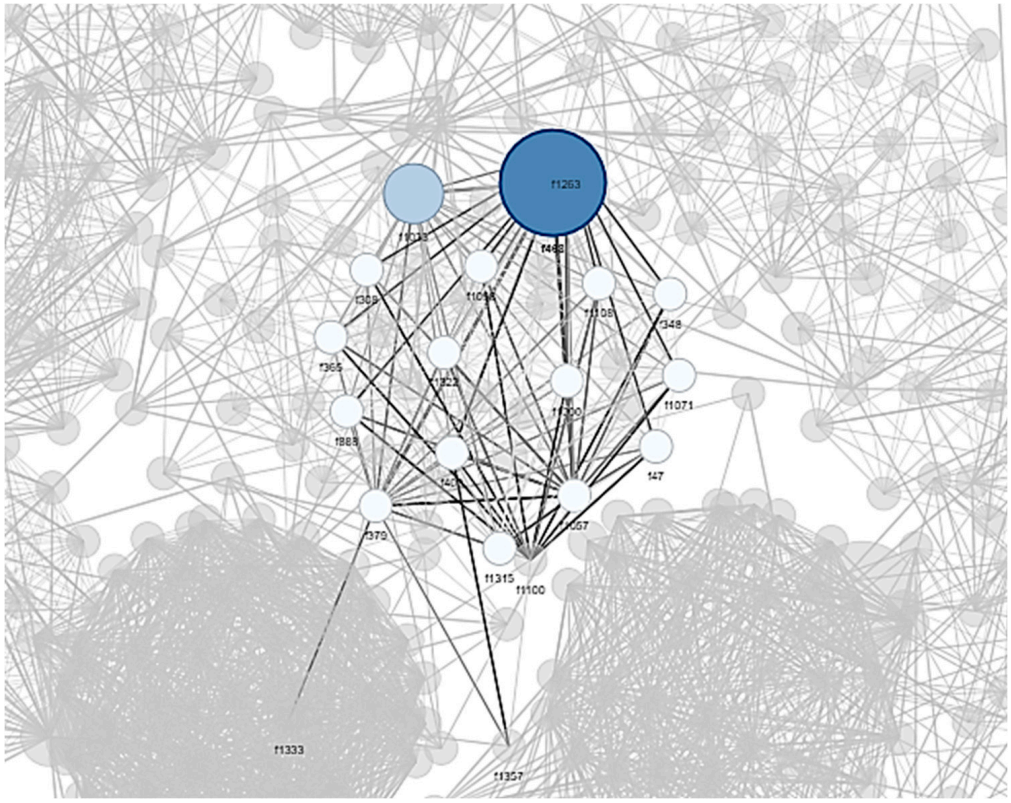


(**a**)



(**b**)

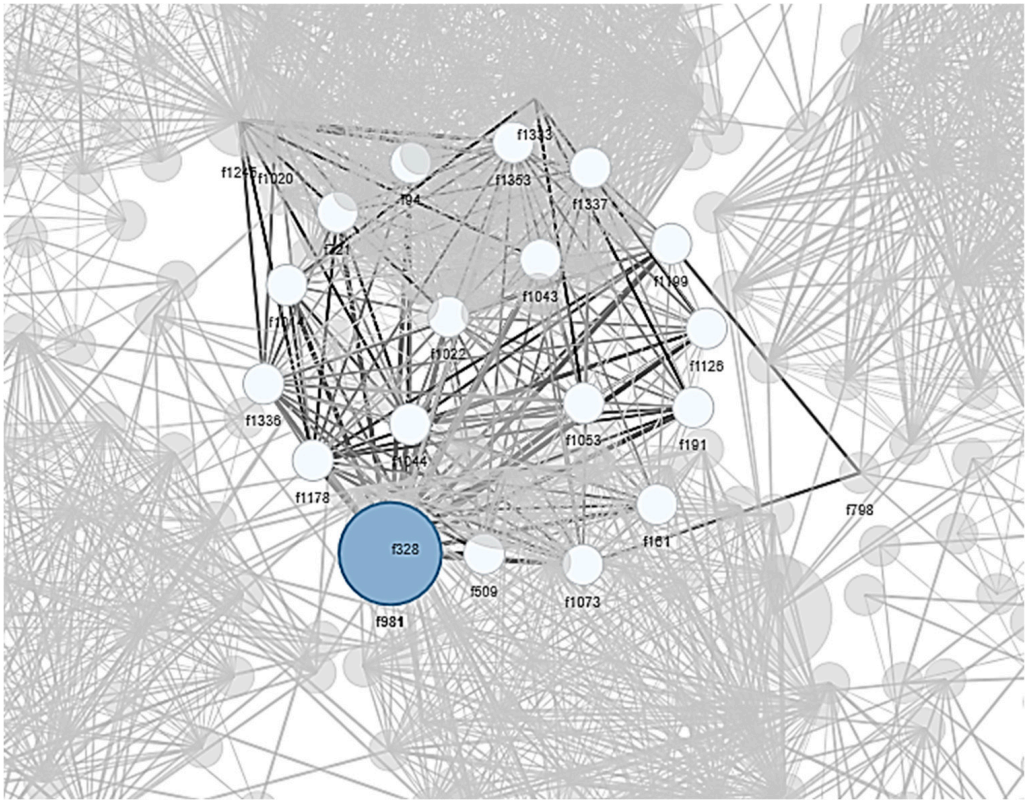
Figure 7. *Cont*.



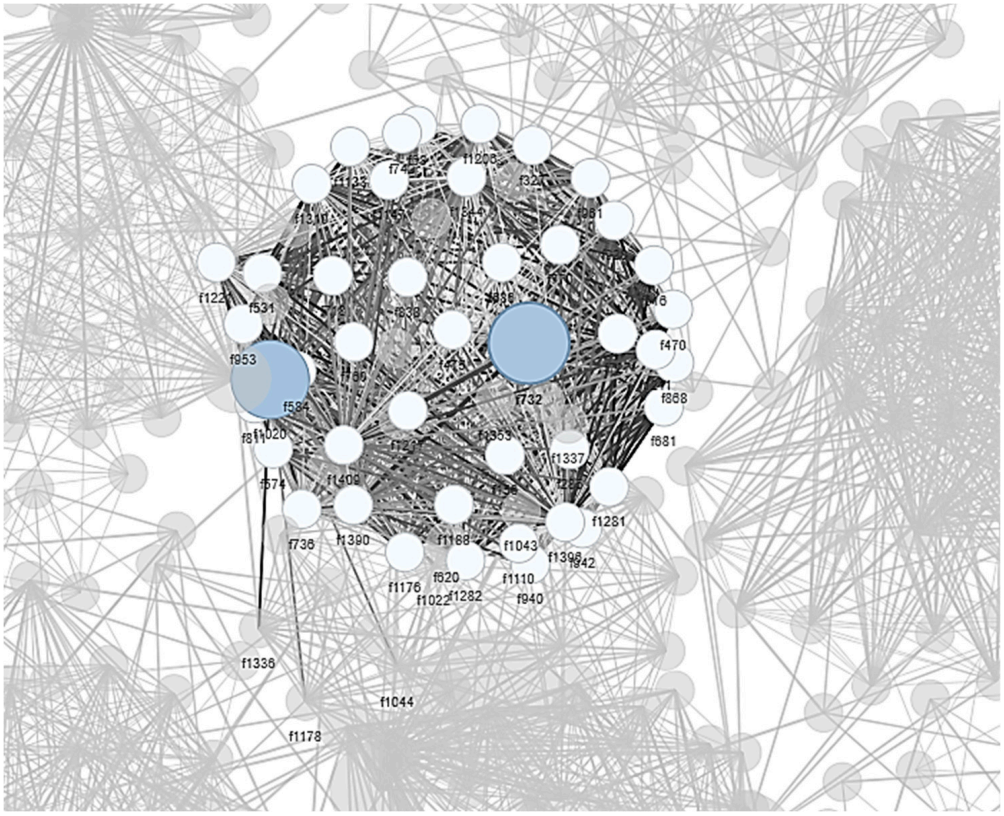
(**c**)

(**d**)

Figure 7. *Cont*.



(**e**)



(**f**)

Figure 7. Local subnetworks for the six highest-ranked nodes according to SL-WLEN: (**a**) f625, (**b**) f1397, (**c**) f468, (**d**) f506, (**e**) f981, and (**f**) f732. Node size represents centrality value, and blue color indicates the analyzed central node.

The identification of these central characteristics through SL-WLEN reveals not only nodes important for final product quality but also their role in manufacturing system stability. The connection structure of these nodes suggests that they are critical points for maintaining process coherence and stability; alterations in these characteristics could propagate extensively through the network due to their multilevel connectivity patterns. This complements the traditional approach based solely on nodal values by considering how these characteristics act as system stabilizers through their interconnections. For example, the dense and cohesive structure of node f625 suggests that it is crucial not only for final quality but also for maintaining the operational stability of the manufacturing process.

6. Robustness Analysis of the SL-WLEN Metric

To evaluate the robustness of the SL-WLEN metric, we adapted the methodology proposed by [42], which continues to be employed in contemporary research, such as in the study [43], who developed a systematic framework to analyze how classical centrality measures (degree, betweenness, closeness, and eigenvector) maintain their consistency under different conditions of error or perturbation in network data. The same perturbation and evaluation techniques were applied to our composite SL-WLEN metric, which, unlike classical metrics, incorporates both structural aspects and nodal values in its calculation. The importance of this analysis lies in that, in real situations, networks may be subject to various types of modifications or errors in their structure.

The process began with selecting a representative sample of the network, balancing computational efficiency and structural representativeness. Given that the complete network consists of 1555 nodes, a robustness analysis on the entire network would be computationally intensive and time-demanding. Therefore, a sample size of 100 nodes was determined, large enough to capture the network structure without compromising analysis viability. To ensure representativeness, stratified sampling based on connectivity distribution was implemented, following the *power-law* distribution observed in real networks. Strata were defined according to node degree, classifying them into high, medium, and low connectivity. Node allocation in each stratum was performed using the formula $n_h = n\left(\frac{N_h}{N}\right)$, where n_h is the sample size for stratum h, n is the total sample size, N_h is the stratum size in the population, and N is the total number of nodes. To ensure balanced network representation, the sample distribution was adjusted, allocating 20% to highly connected nodes (*hubs*), 60% to medium connectivity nodes, and 20% to peripheral nodes. This allowed for capturing the global network structure while optimizing computational resources during test execution.

Four fundamental types of error that can occur in real networks were considered. Node removal simulates scenarios where data from some actors are lost, randomly selecting $n \times p$ nodes, where n is the total number of nodes and p is the modification proportion (0.01, 0.05, etc.). Node addition represents situations where new actors are incorporated into the network, creating $n \times p$ new nodes with degrees similar to randomly selected existing nodes. Edge removal simulates cases where existing connections are lost, randomly selecting $m \times p$ edges, where m is the total number of edges. Edge addition represents scenarios where $m \times p$ new connections are created between previously unconnected nodes.

The selection of perturbation levels was made considering a spectrum ranging from minimal modifications to substantial network changes. Levels of 1% were used, representing minimal errors or minor natural changes in the network; levels of 5% and 10% simulated moderate perturbations; levels of 25% represented significant structure alteration; and levels of 50% simulated extreme network modification. This gradation allows for the evaluation of the metric's sensitivity to small perturbations, its resistance to moderate changes, and its behavior under extreme conditions.

To evaluate different aspects of robustness, five complementary metrics were implemented. The Top 1 metric measures the proportion of times the most important node maintains first position after modification, Top 3 indicates the frequency with which it remains among the top three, and Top 10% represents the proportion of times it stays within the first decile. For example, if a node has values of 0.980 in Top 1, 0.900 in Top 3, and 1.000 in Top 10%, it means that in 98% of replications it retained first position, in 90% it remained among the top three, and in all replications it stayed within the first decile.

The Overlap measure calculates the normalized intersection between the upper deciles of the original and modified networks, defined as $| A \cap B | / | A \cup B |$, where A and B represent the sets of nodes in the first decile of each network. Its value varies between 0 and 1, indicating the degree of coincidence between both networks; for example, a value of 0.720 means that 72% of nodes in the first decile are the same in both versions. On the other hand, R^2 represents the square of the Pearson correlation between the SL-WLEN values of the original and modified network. Values close to 1, such as 0.997, indicate that the relative order of nodes is preserved almost perfectly, while lower values reflect a greater discrepancy in ordering.

The validation process was designed at multiple levels to ensure the robustness of results. Fifty replications were performed, generating in each one a modified version of the network, called "test network", on which modifications were applied. The evaluation included an independent analysis of each combination of error type and level, calculating the five robustness measures and averaging the results. The entire process was documented, recording changes in each replication, ensuring modification traceability, and generating detailed reports.

Tables 5–8 show the robustness results of the SL-WLEN metric under different network perturbation levels.

The test results demonstrate that the SL-WLEN metric exhibits robust and reliable behavior under different network perturbation conditions. The metric shows notable stability against element removal, particularly in the case of edges, where it maintains Top1/3/10% values above 0.900 even with 50% modifications. For node removal, the metric preserves its stability up to 25% modification, with Top1/3/10% values equal to or greater than 0.800, and maintains R^2 values above 0.92 even with 50% alterations.

Table 5. Robustness test results for the SL-WLEN metric under different perturbation levels: node removal case.

Proportion	Top1	Top3	Top10%	Overlap	R^2
0.01	0.980	0.980	0.980	0.687	0.996
0.05	0.920	0.920	0.920	0.636	0.994
0.10	0.900	0.900	0.900	0.580	0.991
0.25	0.800	0.800	0.800	0.381	0.978
0.50	0.500	0.500	0.500	0.213	0.921

Table 6. Robustness test results for the SL-WLEN metric under different perturbation levels: node addition case.

Error Type	Top1	Top3	Top10	Overlap	R^2
0.01	0.960	0.960	0.960	0.684	0.989
0.05	0.840	0.840	0.840	0.533	0.963
0.10	0.640	0.640	0.640	0.441	0.925
0.25	0.440	0.440	0.460	0.275	0.833
0.50	0.480	0.480	0.480	0.175	0.766

Table 7. Robustness test results for the SL-WLEN metric under different perturbation levels: edge removal case.

Error Type	Top1	Top3	Top10	Overlap	R^2
0.01	1.000	1.000	1.000	0.720	0.997
0.05	0.980	0.980	1.000	0.759	0.998
0.10	1.000	1.000	1.000	0.726	0.998
0.25	1.000	1.000	1.000	0.632	0.994
0.50	0.900	0.900	1.000	0.553	0.983

Table 8. Robustness test results for the SL-WLEN metric under different perturbation levels: edge addition case.

Error Type	Top1	Top3	Top10	Overlap	R^2
0.01	0.980	0.980	0.980	0.734	0.994
0.05	0.840	0.840	0.840	0.569	0.967
0.10	0.660	0.660	0.700	0.530	0.928
0.25	0.580	0.580	0.600	0.313	0.862
0.50	0.640	0.640	0.660	0.225	0.797

Regarding element addition, both for nodes and edges, the metric shows progressive deterioration starting from 10% modification, demonstrating greater sensitivity to the incorporation of new elements than to their removal. This sensitivity is reflected in a substantial decrease in consistency for large-scale modifications, where R^2 decreases to approximately 0.77. Nevertheless, the metric maintains high reliability in scenarios with small perturbations of 1–5%.

The practical implications of these results confirm that the SL-WLEN metric is particularly effective in identifying and maintaining the hierarchy of the most important nodes in the network, even under conditions of moderate data loss. Its greater sensitivity to the incorporation of new elements suggests the need for caution when making modifications that exceed 25% of the network structure. These findings validate the robustness and utility of the SL-WLEN metric for complex network analysis, demonstrating its capability to maintain consistency in identifying critical nodes under various perturbation conditions.

7. Performance of the Proposed Metric

In this section, the results of empirical experiments conducted to test the performance of the proposed centrality metric against several real-world networks are presented.

7.1. Test Parameters

For the execution of the experiments, the R programming language (version 4.3.3) was used on the RStudio development platform (version 2024.04.2+764). The computer employed runs Windows 10, with an Intel Core i3 processor up to 3.5 GHz and 16 GB of RAM. The influence of nodes is calculated using each of the selected centrality metrics, and they are then sorted from highest to lowest impact to form a TopK of most influential nodes, with $k = 10$, as proposed by [21,44].

SWLAN consists of four parameters, ξ, ζ, δ, and γ, whose values represent the contribution of each component to the final influence value. These parameters have been established equally at 0.25 following the proposal of [18,21]. Meanwhile, the parameter L, which determines the maximum neighborhood level, has been defined after testing its effectiveness with various values.

7.2. Description of Test Datasets

The experiments were conducted on the following six doubly weighted complex networks (weight on edges and valued nodal attributes) of undirected type: (1) Bitcoin transaction network, (2) Sioux Falls city transport flow network, (3) energy flow network, (4) Reddit social network interaction network, (5) Global Urban Network, and (6) quality control network. The use of real and artificial networks is a commonly viable practice for validating node centrality metrics, allowing for the evaluation of their performance in diverse network structures that are representative of real-world scenarios [45].

The Bitcoin transaction network represents a trust system among users trading with Bitcoin, where each node (5881 in total) symbolizes an individual user with an importance value determined by their activity on the platform, while edges (35,592 connections, 89% positive) indicate trust assessments between users with weights. The data were extracted from the Bitcoin OTC platform over a 2.5-year period (January 2014 to April 2017), constituting the first weighted and signed network available for research on reputation systems in anonymous environments.

The Sioux Falls city transport flow network models an urban traffic flow system where each node (24 zones) symbolizes a specific city intersection with an importance value determined by its strategic position, while edges (76 connections) indicate pathways with flow capacities between 2 and 10 units and a constant length of 0.15. This standardized dataset has been utilized as a benchmark in transportation research for traffic assignment studies and road planning.

The bus energy flow network represents an electrical power system where each node (118 buses) symbolizes a substation with an importance value determined by its voltage level (between 100 and 110 kV), while edges indicate transmission lines between these buses with their respective impedances. The dataset originates from a digitization of the American Midwest electrical system from 1962, converted to standard format to serve as a test case in power flow analysis and electrical network stability studies.

The Reddit social network interaction network represents a system of connections between online communities where each node (55,863 subreddits) symbolizes a thematic forum with an importance value determined by its activity, while edges (858,490 connections) indicate hyperlinks between subreddits with an associated sentiment. The data were extracted from Reddit over a 2.5-year period (2014–2017), identifying hyperlinks in post titles and bodies, with temporal metadata and textual property vectors, as part of a research project on interactions and conflicts between online communities.

The Global Urban Network represents a georeferenced spatial system where each node symbolizes an urban element with an importance value determined by its hosting capacity, while edges indicate physical connections with traffic levels as weights. The data were processed using Euclidean buffers of 100 m for nodal attributes and a threshold of 50 m for edge attributes, available in GeoJSON format for multiple global cities, enabling consistent spatial analyses across diverse urban contexts.

The quality control network for chip manufacturing comprises 1763 product observations with 1555 binary quality features, where nodes represent features with importance values determined through logistic regression with Lasso regularization, and edges show significant statistical correlations between features (measured with the Phi coefficient) above an established significance threshold, allowing for the identification of critical control points in the manufacturing process.

The reader can access the data of the six networks used in this study through the following links: the Bitcoin transaction network at https://snap.stanford.edu/data/soc-sign-bitcoin-otc.html (accessed on 7 January 2025), the Sioux Falls transport network at https://raw.githubusercontent.com/bstabler/TransportationNetworks/master/

SiouxFalls/SiouxFalls_net.tntp (accessed on 7 January 2025), the energy flow network at https://labs.ece.uw.edu/pstca/pf118/ieee118cdf.txt (accessed on 7 January 2025), the Reddit social interaction network at https://snap.stanford.edu/data/soc-RedditHyperlinks.html (accessed on 7 January 2025), the Global Urban Network at https://figshare.com/articles/dataset/Global_Urban_Network_Dataset/22124219?file=40160356 (accessed on 7 January 2025), and the quality control network at https://www.kaggle.com/datasets/arbazkhan971/anomaly-detection (accessed on 19 November 2024).

7.3. Performance Evaluation

The experiments are based on the application of the SIR (Susceptible-Infected-Recovered) model proposed by [46], using Kendall's correlation coefficient as a comparative method [47,48]. This allows for the examination of how information propagates through nodes determined as influential, and it evaluates the efficiency of such diffusion in the network structure [18]. The evaluation compares the effectiveness of various centrality metrics, including SLWLAN through SIR simulations. The TopK nodes identified by each metric function as initial infection seeds, evaluating their propagation with parameters λ and μ. Kendall's τ coefficient quantifies the correlation between the original centrality ranking and the actual diffusion capacity observed in the SIR model, thus validating the predictive accuracy of each metric.

The SIR model is recognized for examining propagation dynamics in complex network systems. Each element is categorized as Susceptible (S), Infected (I), or Removed (R). When considering interactions between elements u and v, the contagion mechanism follows two principles expressed in the following Equations (17) and (18):

$$S(u) + I(v) \xrightarrow{\lambda} I(u) + I(v) \tag{17}$$

$$I(u) \xrightarrow{u} R(u) \tag{18}$$

The parameter λ represents the infection rate while μ indicates the recovery rate. Both equations in combination describe how an element v (already infected) can transform a susceptible element u with probability λ, incorporating it into the infected group.

For the evaluation, each centrality indicator identifies the TopK most relevant elements, which function as initial infection foci in the model. Following the SIR logic, these elements can transmit the infection to their neighboring elements with probability λ or can recover with probability μ. The function $F(t)$ quantifies the sum of both infected and recovered elements at a specific moment t, as shown in Equation (19). This value serves as an indicator of the influence capacity of the initial foci, where t represents a temporal unit in the simulation.

$$F(t) = \frac{N_{I(t)} + N_{R(t)}}{N} \tag{19}$$

The variables $N_{I(t)}$ and $N_{R(t)}$ correspond to the number of elements in infected and recovered states, respectively.

Kendall's τ coefficient is widely used to compare hierarchically ordered elements [6]. In this coefficient, each observation represents the position of a node according to a centrality metric. Let $A = \{a_1, a_2, \ldots, a_2, \ldots, a_N\}$ and $B = \{b_1, b_2, \ldots, b_2, \ldots, b_N\}$ be two lists that order the nodes according to their influence scores. List A is obtained from the proposed centrality metric, while B is generated from the application of SIR. Suppose two pairs of common nodes (a_i, b_i) and (a_j, b_j) belong to sets A and B, respectively. According to Kendall's τ coefficient, if $(a_i - a_j) - (b_i - b_j) > 0$, then the influence scores of these

pairs of nodes are concordant; otherwise, they are considered discordant. Equation (20) formally defines Kendall's τ coefficient, as follows:

$$\tau = \frac{N_c - N_d}{0.5\, N(N-1)} \tag{20}$$

where N_c and N_d represent the number of concordant and discordant pairs, respectively. The coefficient varies between -1 and 1, where higher values indicate a greater similarity between the ranking lists.

7.4. Benchmark Metrics

In this work, the correlation between SL-WLEN and centrality metrics based on topological structures is evaluated, such as semi-local centrality (SC), k-shell decomposition (KS), Local Relative Average Shortest Path (LRASP), weighted hybrid centrality (WHC), and influential node based on Average Shortest Path (INASP). Additionally, the performance of SL-WLEN is compared with centrality metrics based on nodal attributes, such as node-weighted degree centrality (NWDegC), node-weighted harmonic centrality (NWHC), node-weighted betweenness centrality (NWBC), and modified node-weighted eigenvector centrality (MNWEC). The evaluation was conducted using constant configurations for all methods, ensuring an equitable comparison.

7.5. Analysis of Results

7.5.1. Analysis of the L Parameter of SL-WLEN

The maximum exploration level (L) is the sole adjustable parameter defining SL-WLEN's neighborhood scope. This parameter is present in the definitions of DegreeSemiLocal and SemilocalNodeLexOrder of SL-WLEN. The value of L plays an important role in identifying influential nodes in the network, as L is applied as a hop count to extract the local subgraph through the extended neighborhood connectivity (ENC) concept. L is used to determine the scope of the local subgraph.

Traditional semi-local centrality metrics generally only consider first- and second-level neighbors [11]. However, exploring higher values of the L parameter could improve precision levels for identifying influential nodes in networks. Based on this premise, we evaluated various values of L in the DegreeSemiLocal and SemilocalNodeLexOrder components to optimize the performance of SL-WLEN. Table 9 presents the comparative evaluation between various network structures using Kendall's τ coefficient. The data presented correspond to simulations performed with fixed recovery $(\mu = 1)$ and infection $(\lambda = 0.1)$ parameters.

Table 9. Evaluation of SL-WLEN with different levels of neighborhoods through the L parameter.

Network	L = 1	L = 2	L = 3	L = 4	L = 5
Bitcoin transactions	0.672	0.723	0.674	0.689	0.666
Sioux Falls City transport flow	0.632	0.689	**0.798**	0.654	0.632
Energy flow network	0.647	0.656	**0.753**	**0.78**	0.644
Reddit social network interactions	0.659	**0.783**	0.632	0.632	0.633
Global urban network	0.654	**0.732**	0.654	0.635	0.639
Quality control network	0.623	**0.771**	0.639	0.649	0.697
Average	0.648	0.726	0.692	0.673	0.652

In the table, the optimal values appear highlighted in bold to facilitate their identification. Upon examining the averages row, it is evident that the SL-WLEN metric achieves its maximum effectiveness when configured with a neighborhood level $L = 2$. Consequently,

the following experimental phases with the SL-WLEN metric focus specifically on this neighborhood level.

7.5.2. Comparison Based on Kendall's τ Coefficient

Kendall's τ coefficient functions as a standard criterion when evaluating centrality metrics in the context of the SIR model. The experiments were conducted maintaining $\mu = 1$ while modifying λ in the spectrum from 0.01 to 0.1, in accordance with the work of [44]. Figure 8 illustrates the results of Kendall's τ coefficient, comparing the SL-WLEN metric with centrality metrics based on connectivity (KS, CS, LRASP, WHC, INASP) across the test networks. Meanwhile, Figure 9 presents the comparison between SL-WLEN and various metrics based on nodal properties (NW-DegC, NW-BC, MW-BC, NW-BC, MW-WC). This analysis encompasses six distinct networks: Bitcoin transaction network, Sioux Falls City transportation flow, energy flow network, Reddit social network interactions, global urban network, and quality control network.

The experiment examines the correlation between propagation dynamics and various centrality metrics under different values of λ. The parameter $F(t)$ quantifies the total of both infected and removed nodes at time t, naturally increasing with temporal advancement. In the presented graphs, the rank correlation r is equivalent to Kendall's coefficient, reflecting the relationship between the accumulation of infected nodes according to each metric and the $F(t)$ value. Higher correlation values indicate greater predictive capacity regarding propagation potential.

The simulation results demonstrate that the SL-WLEN metric maintains consistently high values of the τ coefficient across all analyzed networks. According to the graphs presented in Figures 8 and 9, it can be observed that SL-WLEN improves Kendall's τ coefficient compared to both connectivity-based metrics and nodal attribute-based metrics. Specifically, in the Bitcoin transaction network, SL-WLEN improves between 8% and 10% over SC; in the energy flow network, where SL-WLEN maintains a particularly notable advantage, the improvement ranges between 15% and 18% compared to traditional metrics; in the Global Urban Network, improvement is between 7% and 9% over NW-DegC; and in the quality control network, it is between 12% and 15% over NW-BC.

The variability in the performance of the SL-WLEN method across the analyzed networks can be attributed to the inherent heterogeneity in the joint distribution of edge weights and nodal attributes. Each network exhibits a distinct structural correlation between these components, where SL-WLEN effectively captures this multidimensional complexity. The variations observed in Kendall's τ coefficient demonstrate the sensitivity and adaptability of the method to diverse network topologies. This phenomenon corroborates the robustness and general utility of SL-WLEN as an effective analytical tool for centrality evaluation in complex networks with heterogeneous topological characteristics.

Specifically, SL-WLEN demonstrates superior performance when edge weights and nodal values provide complementary information about node importance in propagation dynamics; conversely, when these signals exhibit redundancy or conflicting patterns—as can occur in networks with specific physical constraints, such as in the Sioux Falls transport flow network—comparative performance fluctuations are observed.

A deeper analysis reveals that these variations can be explained through the following interrelated factors: First, in the energy flow network and quality control network, cascade propagation phenomena align closely with the underlying assumptions of the SIR model, whereas in the Sioux Falls transport flow network, the deterministic nature of flow limits the advantages of the probabilistic approach. Second, in networks such as the Bitcoin transaction network, the asymmetric trust structure allows the SemilocalNodeLexOrder component to effectively preserve the influence of high-reputation nodes, while in the

Sioux Falls transport flow network, structural predetermination reduces the added value of lexicographic ordering. Third, the degree of modular structure affects relative performance, as observed in the Reddit social interaction network, where the pronounced community structure generates propagation dynamics that vary according to the λ parameter of the SIR model.

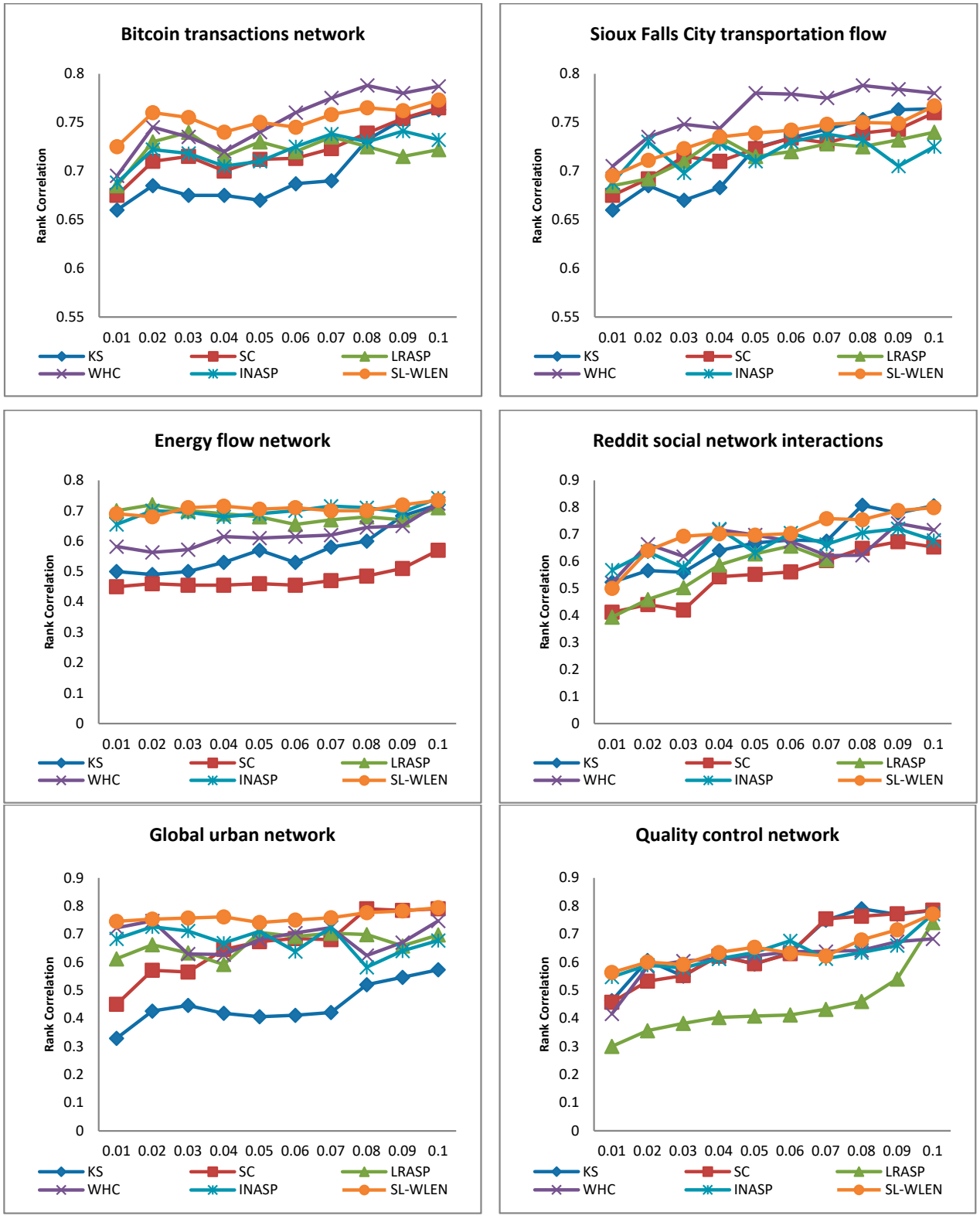


Figure 8. Kendall's τ coefficient results for SL-WLEN and connectivity-based centrality metrics across different test networks.

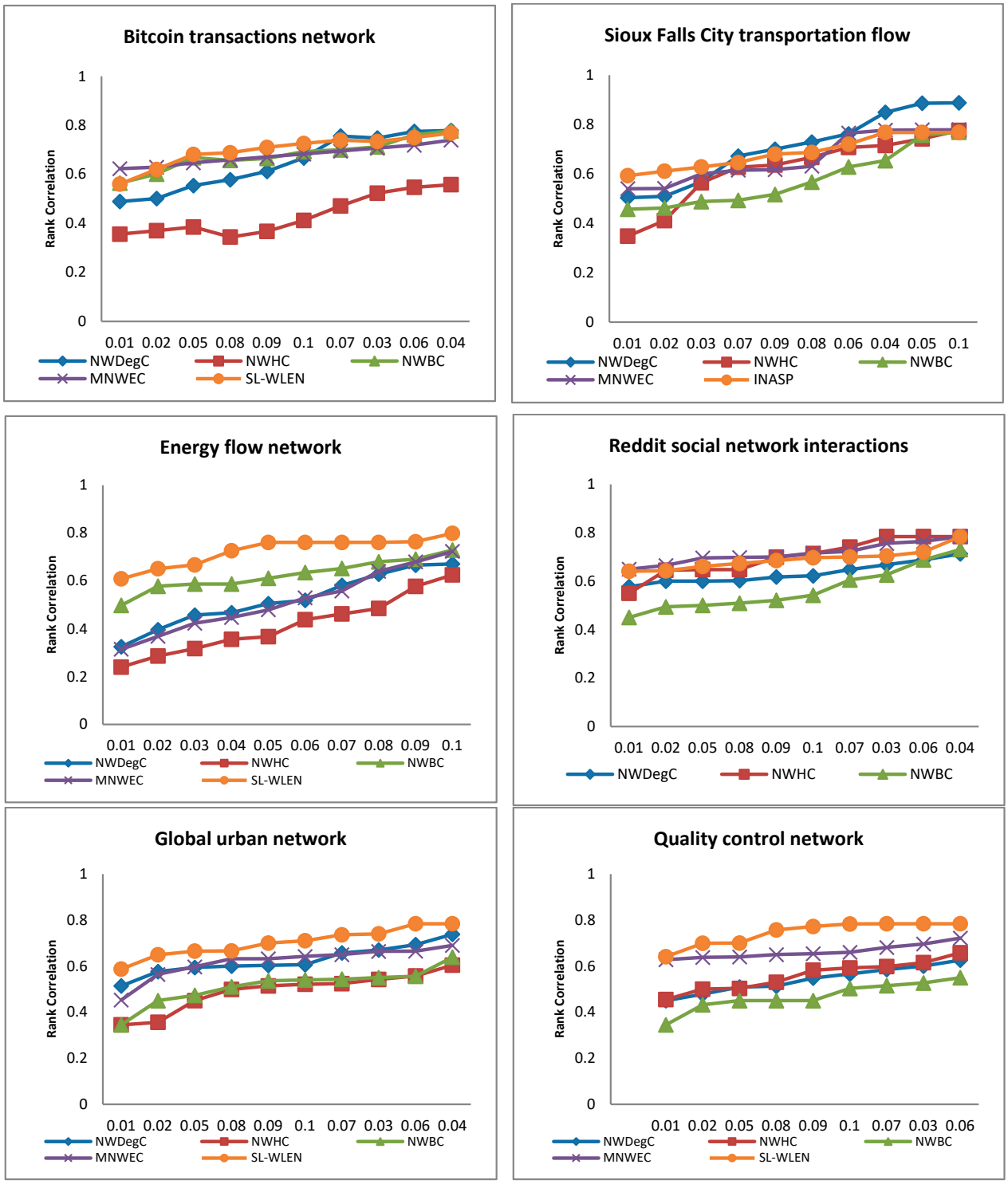


Figure 9. Kendall's τ coefficient results for SL-WLEN and nodal attribute-based centrality metrics across different test networks.

The analysis also identifies specific limitations. When there is a high correlation between topological importance and nodal values (as in the Sioux Falls network), the components of SL-WLEN may become redundant; performance varies considerably with different values of parameter L, requiring specific calibration; lexicographic ordering introduces additional computational cost that may limit applicability in extremely dense networks; and in systems where spatial or physical constraints rigorously determine structure, the added value of the lexicographic approach decreases considerably.

This performance variability demonstrates that, while SL-WLEN generally provides superior results, its effectiveness is conditioned by specific network characteristics, allowing for the identification of scenarios where its application is optimal as well as those where alternative methods might be more appropriate.

7.6. Complexity Analysis

The complexity of SL-WLEN involves calculating local and semi-local influence. Local influence (DegreeLocal) has complexity $O(1)$, as it only requires dividing the node degree by the size of the network. Semi-local influence (DegreeSemiLocal) has complexity $O(N_v + E_v\, logN_v)$, as it depends on subgraph extraction through ENC $O(N_v)$ and processing neighbors up to level L.

The calculation of SemilocalNodeLexOrder adds the complexity of the lexicographic ordering of neighbors, which at each level l is $O(N_v\ log\ N_v)$. The partial contributions by level have complexity $O(N_v)$, and their accumulation up to L is $O(L)$.

Together, the total complexity of SL-WLEN is $O(L{\cdot}N_v\ log\ N_v + E_v\ logN_v)$, dominated by lexicographic ordering. This additional cost allows for better capturing the relative importance of nodes through their attributes, differentiating it from other semi-local metrics. This is consistent with what is noted by [49], who argue that not all nodes have the same level of importance within a network, depending on their impact on the structure and dynamic behavior of the system.

8. Discussion

The SL-WLEN metric demonstrates applicability in the test networks employed in this study. Results across six heterogeneous networks (Bitcoin, Sioux Falls, energy flow, Reddit, global urban, and quality control) evidence its versatility and robustness in different contexts. In chip manufacturing environments, SL-WLEN enables the identification of critical features that both directly and indirectly affect the quality of the final product. The capacity to integrate nodal value and topological structure facilitates the discovery of strategic control points in the production process, proving particularly valuable for improving inspection processes, designing predictive monitoring systems, and prioritizing investments in quality improvements. Nodes f625 and f1397, identified as the most central according to SL-WLEN, are not only individually important but act as system stabilizers due to their multilevel connectivity patterns. This validates the utility of the method for capturing critical elements in the manufacturing context.

Beyond the principal case study, SL-WLEN demonstrates applicative potential in financial and trust networks, as evidenced in the Bitcoin network, where it identifies key users in reputation systems with 8–10% greater precision than traditional metrics. Its applicability extends to transportation systems, characterizing strategic intersections where interventions would have a greater impact on global flow; energy networks, identifying critical substations for system stability with a 15–18% improvement over conventional metrics; social networks, detecting communities with greater information diffusion power; and urban planning, identifying strategic locations for infrastructure development.

The comparative analysis using Kendall's τ coefficient reveals that SL-WLEN consistently outperforms metrics based solely on connectivity (SC, KS, LRASP, WHC, INASP) or exclusively on nodal attributes (NWDegC, NWHC, NWBC, MNWEC). This superiority is due to its effective integration of structure and attributes, capturing their interaction in a manner that better reflects the reality of complex systems where importance depends on both factors jointly. The lexicographic ordering presents a distinctive advantage by preserving the relative hierarchy of important nodes, avoiding the dilution of their influence that typically occurs with methods based on statistical aggregations. Additionally,

SL-WLEN achieves a favorable computational balance, maintaining acceptable complexity, which offers greater precision than local metrics without incurring the prohibitive cost characteristic of global metrics for large-scale networks.

Despite its advantages, SL-WLEN presents important limitations that must be considered for its optimal application. This metric shows sensitivity to the joint distribution of edge weights and nodal attributes, exhibiting performance variations in networks where these components present redundancy or conflicting patterns, as observed in certain fluctuations in the Sioux Falls network. The effectiveness of this method is conditioned by the selection of the neighborhood level L, and, although optimal on average for the analyzed networks, each network type might require a specific value, implying a prior calibration process. Robustness analyses reveal an asymmetry in SL-WLEN's response to structural modifications. It maintains greater consistency when elements are removed than when they are added, suggesting potential challenges in application to dynamic networks with rapid growth, where it might require more frequent recalculation to maintain its precision.

From a practical perspective, although SL-WLEN is computationally more efficient than global metrics, the lexicographic ordering adds an additional cost that could be a consideration in extremely large networks. For systems with millions of nodes, this incremental cost could compromise the practical applicability of the method in analyses requiring frequent or real-time updates. The effectiveness of SL-WLEN also depends on the adequate normalization of nodal values to allow meaningful comparisons, implying prior data processing that could be complex in cases where attributes present highly skewed distributions or significant outliers. While conceptually adaptable to different typologies, the present study focused primarily on undirected networks, so its application in directed networks might require methodological modifications to adequately capture the characteristic asymmetry of these relationships. The component weighting parameters (α and β) were equitably established following previous proposals, but different network types might benefit from alternative configurations, with the identification of optimal values specific to each domain being an additional practical challenge.

In terms of scalability, the current implementation shows performance deterioration in scenarios with substantial network modifications (greater than 25% in additions), suggesting potential limitations for its application in systems with high structural dynamics or requiring real-time analysis. To address these limitations, future research could explore the development of adaptive versions that automatically adjust parameters according to specific characteristics of each network; incremental implementations that allow updating centrality values without complete recalculations when the network experiences localized changes; parallelization strategies to improve performance in large-scale networks; extensions for temporal networks that capture the dynamic evolution of centrality; and integrations with machine learning techniques to predict changes in the relative importance of nodes. In conclusion, while SL-WLEN represents an advance in the identification of influential nodes in complex networks with weighted edges and nodal attributes, its application requires careful consideration of its limitations and the specific implementation context, balancing precision and computational efficiency according to the particular requirements of each application domain.

9. Conclusions

This paper proposes SL-WLEN as a weighted semi-local centrality metric based on the integration of lexicographic ordering and the extended neighborhood concept for identifying influential nodes in complex quality control networks. Beyond node importance, SL-WLEN incorporates both topological structure and nodal values in its evaluation, considering the following four main components: local influence by connectivity, local node

influence, semi-local degree influence, and semi-local influence based on lexicographic ordering. By applying a distributed approach that analyzes subgraphs per node and utilizing lexicographic ordering to evaluate hierarchical importance, SL-WLEN provides an effective balance between accuracy and computational complexity.

The simulation results demonstrate that the SL-WLEN metric maintains consistently high values of Kendall's τ coefficient across all analyzed networks, outperforming both connectivity-based metrics and nodal attribute-based metrics. Specifically, SL-WLEN improves between 8% and 10% over SC in the Bitcoin transactions network, between 15% and 18% in the energy flow network, between 7% and 9% over NW-DegC in the global urban network, and between 12% and 15% over NW-BC in the quality control network. Additionally, the analysis of the L parameter demonstrates that the metric achieves its maximum effectiveness when configured with a neighborhood level $L = 2$.

Numerical robustness tests demonstrate SL-WLEN's high stability, especially against element removal, maintaining its consistency even with significant network modifications. However, the metric shows greater sensitivity to the incorporation of new elements, suggesting areas for improvement in future work. The extension of the metric to consider network temporal dynamics and its adaptation for different types of complex networks represents a promising direction for subsequent research.

Author Contributions: Conceptualization, M.F.O.M.; methodology, M.F.O.M. and R.I.Y.M.; software, M.F.O.M.; validation, M.F.O.M. and R.I.Y.M.; formal analysis, M.F.O.M.; investigation, M.F.O.M.; resources, M.F.O.M.; data curation, M.F.O.M.; writing—original draft preparation, M.F.O.M.; writing—review and editing, M.F.O.M. and R.I.Y.M.; visualization, M.F.O.M.; supervision, M.F.O.M.; project administration, M.F.O.M.; funding acquisition, M.F.O.M. All authors have read and agreed to the published version of the manuscript.

Funding: The Article Processing Charges (APC) were funded by Pontifical Catholic University of Ecuador Ibarra.

Data Availability Statement: The original data presented in the study are openly available in Kaggle at https://www.kaggle.com/datasets/arbazkhan971/anomaly-detection, accessed on 19 November 2023.

Acknowledgments: This research was conducted within the framework of the Mathematics Ph.D. Program at the Instituto de Estudios Superiores de Investigación y Postgrado de Venezuela, to which we express our gratitude.

Conflicts of Interest: The authors declare no conflicts of interest in this work.

References

1. Liu, Z.; Gu, Z.; Liu, P. Artificial neural networks. In *Transportation Big Data: Theory and Methods*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 347–382. [CrossRef]
2. Cuevas, E.; Avila, K.; Islas Toski, M.; Escobar, H. Complex networks and graph theory. In *Agent-Based Models with MATLAB*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 75–103. [CrossRef]
3. Lazzaro, I.; Milano, M. Network centralities and node ranking. In *Reference Module in Life Sciences*; Elsevier: Amsterdam, The Netherlands, 2024. [CrossRef]
4. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]
5. Brin, S.; Page, L. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [CrossRef]
6. Sheng, J.; Dai, J.; Wang, B.; Duan, G.; Long, J.; Zhang, J.; Guan, W. Identifying influential nodes in complex networks based on global and local structure. *Phys. A Stat. Mech. Appl.* **2020**, *541*, 123262. [CrossRef]
7. Abbasi, A.; Hossain, L. Hybrid centrality measures for binary and weighted networks. In *Complex Networks*; Springer: Berlin, Germany, 2013; pp. 1–7.

8. Benyahia, O.; Largeron, C. Centrality for graphs with numerical attributes. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Paris, France, 25–28 August 2015; pp. 1348–1353. [CrossRef]
9. Kitsak, M.; Gallos, L.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H.E.; Makse, H.A. Identification of influential spreaders in complex networks. *Nat. Phys.* **2010**, *6*, 888–893. [CrossRef]
10. Zeng, A.; Zhang, C.J. Ranking spreaders by decomposing complex networks. *Phys. Lett. A* **2013**, *377*, 1031–1035. [CrossRef]
11. Chen, D.; Lü, L.; Shang, M.S.; Zhang, Y.C.; Zhou, T. Identifying influential nodes in complex networks. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 1777–1787. [CrossRef]
12. Gao, S.; Ma, J.; Chen, Z.; Wang, G.; Xing, C. Ranking the spreading ability of nodes in complex networks based on local structure. *Phys. A Stat. Mech. Appl.* **2014**, *403*, 130–147. [CrossRef]
13. Liu, J.; Xiong, Q.; Shi, W.; Shi, X.; Wang, K. Evaluating the importance of nodes in complex networks. *Phys. A Stat. Mech. Appl.* **2016**, *452*, 209–219. [CrossRef]
14. Opsahl, T.; Agneessens, F.; Skvoretz, J. Node centrality in weighted networks: Generalizing degree and shortest paths. *Soc. Netw.* **2010**, *32*, 245–251. [CrossRef]
15. Puzis, R.; Elovici, Y.; Zilberman, P.; Dolev, S.; Brandes, U. Topology manipulations for speeding betweenness centrality computation. *J. Complex Netw.* **2014**, *3*, 84–112. [CrossRef]
16. Akanmu, A.A.; Wang, F.Z.; Yamoah, F.A. Clique structure and node-weighted centrality measures to predict distribution centre location in the supply chain management. In Proceedings of the Science and Information Conference (SAI), London, UK, 27–29 August 2014; pp. 100–111.
17. Bavelas, A. Communication patterns in task-oriented groups. *J. Acoust. Soc. Am.* **1950**, *22*, 725–730. [CrossRef]
18. Zhang, K.; Zhou, Y.; Long, H.; Wang, C.; Hong, H.; Armaghan, S.M. Towards identifying influential nodes in complex networks using semi-local centrality metrics. *J. King Saud Univ. Comput. Inf. Sci.* **2023**, *35*, 101798. [CrossRef]
19. Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.* **1972**, *2*, 113–120. [CrossRef]
20. Hajarathaiah, K.; Enduri, M.K.; Anamalamudi, S. Efficient algorithm for finding the influential nodes using local relative change of average shortest path. *Phys. A Stat. Mech. Appl.* **2022**, *591*, 126708. [CrossRef]
21. Xiao, Y.; Chen, Y.; Zhang, H.; Zhu, X.; Yang, Y.; Zhu, X. A new semi-local centrality for identifying influential nodes based on local average shortest path with extended neighborhood. *Artif. Intell. Rev.* **2024**, *57*, 115. [CrossRef]
22. Lv, Z.; Zhao, N.; Xiong, F.; Chen, N. A novel measure of identifying influential nodes in complex networks. *Phys. A Stat. Mech. Appl.* **2019**, *523*, 488–497. [CrossRef]
23. Zhao, X.; Liu, F.; Wang, J.; Li, T. Evaluating influential nodes in social networks by local centrality with a coefficient. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 35. [CrossRef]
24. Tripathy, A.; Green, O. Scaling betweenness centrality in dynamic graphs. In Proceedings of the IEEE High Performance Extreme Computing Conference (HPEC), Waltham, MA, USA, 25–27 September 2018; pp. 1–7. [CrossRef]
25. Hajarathaiah, K.; Enduri, M.K.; Dhuli, S.; Anamalamudi, S.; Cenkeramaddi, L.R. Generalization of relative change in a centrality measure to identify vital nodes in complex networks. *IEEE Access* **2023**, *11*, 653–666. [CrossRef]
26. Qu, D.; Liang, W.; Zhang, Y.; Gu, C.; Zhou, G.; Zhan, Y. Research on identification of critical quality features of machining processes based on complex networks and entropy-CRITIC methods. *Computers* **2024**, *13*, 216. [CrossRef]
27. Kulcsár, E.; Gyurika, I.G.; Csiszér, T. Network-based—Quality Function Deployment (NB-QFD): The combination of traditional QFD with network science approach and techniques. *Comput. Ind.* **2021**, *133*, 103592. [CrossRef]
28. Niu, Q.; Cheng, S.; Qiu, Z. Algorithm for correlation diagnosis in multivariate process quality based on the optimal typical correlated component pair group. *Processes* **2024**, *12*, 652. [CrossRef]
29. Charitou, T.; Lallas, E.; Gerogiannis, V.C.; Karageorgos, A. A network modeling and analysis approach for pharma industry regulatory assessment. *IEEE Access* **2024**, *12*, 46470. [CrossRef]
30. Wang, L.; He, F.; Zhao, Y.; Wang, J.; Lu, P.; Ou, Z.; Jia, Y. Complex network-based analysis of inter-basin water transfer networks. *Ecol. Indic.* **2023**, *156*, 111197. [CrossRef]
31. Unicomb, S.; Iñiguez, G.; Gleeson, J.P.; Karsai, M. Dynamics of cascades on burstiness-controlled temporal networks. *Nat. Commun.* **2021**, *12*, 133. [CrossRef]
32. Aleandri, M.; Fragnelli, V.; Moretti, S. Lexicographic ranking based on minimal winning coalitions. *arXiv* **2022**, arXiv:2207.07302. [CrossRef]
33. Ding, J.; Liu, C.; Zheng, Y.; Zhang, Y.; Yu, Z.; Li, R.; Chen, H.; Piao, J.; Wang, H.; Liu, J.; et al. Artificial Intelligence for Complex Network: Potential, Methodology and Application. *arXiv* **2024**, arXiv:2402.16887. [CrossRef]
34. Liu, X. Special Issue "Artificial Intelligence in Complex Networks". *Appl. Sci.* **2024**, *14*, 2822. [CrossRef]
35. Freeman, L.C. A centrality in social networks conceptual clarification. *Soc. Netw.* **2002**, *1*, 238–263. [CrossRef]

36. Sherly, K.D.; Bhattacharjee, S. A weighted hybrid centrality for identifying influential individuals in contact networks. In Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 15–17 July 2022; pp. 1–6.
37. Ma, L.; Ma, C.; Zhang, H.-F.; Wang, B.-H. Identifying influential spreaders in complex networks based on gravity formula. *Phys. A* **2016**, *451*, 205–212. [CrossRef]
38. Zhang, W. Screening node attributes that significantly influence node centrality in the network. *Selforganizology* **2016**, *3*, 114–124.
39. Wei, B.; Liu, J.; Wei, D.; Gao, C.; Deng, Y. Weighted k-shell decomposition for complex networks based on potential edge weights. *Phys. A* **2015**, *420*, 277–283. [CrossRef]
40. Maji, G.; Namtirtha, A.; Dutta, A.; Curado Malta, M. Influential spreaders identification in complex networks with improved k-shell hybrid method. *Expert Syst. Appl.* **2020**, *144*, 113092. [CrossRef]
41. Borgatti, S.P.; Carley, K.M.; Krackhardt, D. On the robustness of centrality measures under conditions of imperfect data. *Soc. Netw.* **2006**, *28*, 124–136. [CrossRef]
42. Everett, M.G.; Borgatti, S.P. Categorical attribute-based centrality: E-I and G-F centrality. *Soc. Netw.* **2012**, *34*, 562–569. [CrossRef]
43. Gnielka, F.M.; Reichel, R.; Blokland, A.; Daser, A.; de Boer, M.; Gannon, C.; Schmidt, A.F.; Schäfer, T.; Huikuri, S.; Staciwa, K.; et al. Missing the mark? Identifying child sexual abuse material forum structure and key-players based on public replies and private messaging networks. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 1459. [CrossRef]
44. Ullah, A.; Wang, B.; Sheng, J.; Sun, X. Escape velocity centrality: Escape influence-based key nodes identification in complex networks. *Appl. Intell.* **2022**, *52*, 16586–16604. [CrossRef]
45. Sun, Z.; Sun, Y.; Chang, X.; Wang, F.; Wang, Q.; Ullah, A.; Shao, J. Finding critical nodes in a complex network from information diffusion and Matthew effect aggregation. *Expert Syst. Appl.* **2023**, *233*, 120927. [CrossRef]
46. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **1927**, *115*, 700–721. [CrossRef]
47. Shao, Z.; Liu, S.; Zhao, Y.; Zhang, L.; Xie, C. Identificación de nodos influyentes en redes complejas basada en vecinos y bordes. *Peer-to-Peer Netw. Appl.* **2019**, *12*, 1528–1537. [CrossRef]
48. Huang, S.; Zong, G.; Zhao, N.; Zhao, X.; Ahmad, A.M. Performance recovery-based fuzzy robust control of networked nonlinear systems against actuator fault: A deferred actuator-switching method. *Fuzzy Sets Syst.* **2024**, *480*, 108858. [CrossRef]
49. Amani, A.M.; Fiol, M.A.; Jalili, M.; Chen, G.; Yu, X.; Stone, L. Measuring the Effect of Node Removal on Laplacian Eigenvalues in Complex Networks. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2023**, *70*, 4146–4158. [CrossRef]

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