INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 21, Issue: 2, Month: April, Year: 2026
Article Number: 7108, https://doi.org/10.15837/ijccc.2026.2.7108

CCC Publications

A Passive Clustering-Based Approach for Important Node Mining in Multi-Relational Networks

H.J. Huang

Haijun Huang
School of Computer Engineering
Jiangsu University of Technology, China
213001, Changzhou, Jiangsu
hjhuang@jsut.edu.cn

Abstract

The identification of key nodes within multi-relational networks presents significant challenges due to the heterogeneity of node attributes and the varying significance of different attributes across distinct relationships. Conventional methods often fail to effectively capture these complexities, leading to suboptimal mining outcomes. To address this issue, a passive clustering-based approach is introduced to enhance the identification of important nodes in multi-relational networks. By constructing an adjacency matrix framework, the network structure is systematically represented, encapsulating the connectivity relationships among nodes. The comprehensive centrality of entity nodes is then evaluated to preliminarily select candidates with substantial network influence. Subsequently, a passive clustering algorithm is applied to categorize nodes into clusters based on attribute similarities, enabling a refined analysis within each cluster. The principle of node centrality metrics is further adapted to assess node importance within and across clusters, thereby mitigating the impact of attribute heterogeneity. Nodes exhibiting weak intra-cluster associations are eliminated, ensuring the robustness of the clustering process. The proposed method demonstrates superior efficiency and scalability, requiring a memory footprint below 160 KB. Furthermore, the computational efficiency of node degree centrality, median centrality, and proximity centrality is improved, with relative computational time ratios of 14.2%, 8.9%, and 8.6%, respectively. These results indicate that the proposed approach effectively captures complex dynamic interactions within undirected and unprivileged multi-relational networks, offering a scalable and computationally efficient solution for important node mining.

Keywords: Multi-relational networks, centrality measures, important node identification, passive clustering, entity nodes.

1 Introduction

In the interdisciplinary field of digital economy and complex systems science, multi relational networks have become a key tool for revealing social dynamics and optimizing resource allocation. With the deep popularization of 5G technology and the rapid development of industrial Internet, the interaction between network nodes presents multi-dimensional and dynamic characteristics. In a multi relational network, individuals, organizations, things, and other relational entities mainly establish various types of relationships based on their own attributes and external environmental factors, including but not limited to cooperation, competition, membership, social interaction, etc. [1, 2]. After the establishment of multiple types of relationships between relationship entities, information, resources, influence, and other factors reflect the dynamic attributes of the multi relationship network in the actual operation and interaction process. They flow in an orderly manner along the relationship path in the network, and the flow direction and flow attributes are directly related to the relationship attributes and strength. Based on the establishment of multiple types of relationships and attribute flow, the network structure continues to dynamically evolve and achieve the main functions of resource aggregation, sharing, and optimized configuration in professional scenarios[3]. In the context of the continuous development of complex network analysis techniques, nodes not only serve as basic units at the micro level of society, but also develop into important perspectives for progressive cognition of complex systems. For example, in the context of intelligent manufacturing, device nodes simultaneously involve production process collaboration, energy consumption monitoring, and supply chain data interaction. This multi relationship coupling characteristic poses a serious challenge to traditional single relationship network analysis methods. Therefore, effective mining of important nodes in complex networks has become one of the important research problems in the fields of computer science and data mining. Although existing research has made progress in the field of node importance assessment, it mostly focuses on static single-layer networks and lacks systematic exploration of node interaction mechanisms in dynamic multi relationship scenarios.

2 Journals reviewed

Several researchers have studied methods for mining important nodes in networks, such as the trust mining algorithm for beacon nodes in large-scale network environments studied in reference [4]. Firstly, based on the distance error evaluation and probability function of beacon nodes in large-scale network environments, the direct trustworthiness of beacon nodes is obtained. Transform trust into influence, use percolation theory to mine the influence of beacon nodes, and determine the beacon nodes with the greatest impact in large-scale network environments. Then, based on the influence of the nodes, the Received Signal Strength Indicator (RSSI) is used to optimize the traditional Distance Vector Hop (DV hop) node localization algorithm. This effect weights the average hop distance of beacon nodes. The influence weight of beacon nodes defines the average hop distance of unknown nodes. The average hop distance information of unknown nodes is taken from beacon nodes with significant impact, which solves the problem of significant positioning errors caused by uncertainty of location targets. Finally, the security status of nodes is reflected based on their level of trust in beacon nodes. The experimental results show that when the number of beacon nodes and communication distance change, the algorithm can accurately locate other nodes in the wide area network environment, and the trust level of the mined nodes can accurately reflect the security status of the nodes. However, when mining the influence of beacon nodes based on percolation theory, the algorithm assumes that all nodes have the same invasion probability, ignoring the heterogeneity of nodes in the actual network.

Reference [5] proposes a key node mining algorithm for air quality systems based on network structure and pollutant transport characteristics, aiming to provide guidance for resource investment. Firstly, establish an air quality network and analyze its structural characteristics. Secondly, based on the diffusion and attenuation mechanisms of air pollutants in the network, a bidirectional transmission key node mining algorithm is proposed, which considers both the inbound and outbound links. Thirdly, a dynamic independent threshold propagation model in a directed weighted network is proposed, and the number of activated nodes is used as the evaluation criterion for key node mining results. The experimental results show that the bidirectional transmission key node mining algorithm can obtain accurate results. However, the state independent threshold propagation model may ignore changes in real-time network conditions, leading to a disconnect between the results of key node mining and the actual network state.

Reference [6] proposed an out of degree graph clustering algorithm (OGC algorithm) for dynamically selecting out of degree boundaries to optimize the clustering range. On this basis, we propose an adaptive seed node mining algorithm based on outdegree (ASMO algorithm). The experiment shows that our algorithm maintains a balance between the cost and benefit of seed node mining, greatly reducing the running time of seed node mining. However, the OGC algorithm is based on a static network topology when dynamically selecting degree boundaries, which cannot effectively cope with node joining/leaving or link changes, resulting in the failure of clustering range optimization.

Reference [7] developed a distributed data mining (DDM) method that utilizes deep learning techniques to distribute the data mining process across the network, reducing the burden on individual nodes and minimizing energy consumption to the greatest extent possible. This study provides a deep learning based DDM method to improve the energy efficiency and load balancing of WSN fusion centers. The DDM model is implemented using adaptive cascaded residual long short-term memory (ACas ResLSTM), which helps integrate the network into multiple layers and place nodes accordingly. The hybrid algorithm of Coati based on defined random numbers and Dolphin Swarm Optimization (DRN-CDSO) was effectively used to optimize the parameters of ResLSTM. Experimental verification shows that the energy efficiency of the proposed technology is 14.5. However, when deploying the ACas ResLSTM model in WSN, the high complexity of the model leads to excessive consumption of computing and storage resources, which goes against the original intention of distributed data mining to reduce node burden.

Literature [4], based on a single attribute or a global unified model, does not fully consider the diversity of node attributes in multiple relational networks, resulting in the deviation of mining results from the actual importance; References [5] and [6] rely on static network assumptions and are difficult to adapt to dynamic node joining/exiting or link changes; The high complexity and global computation strategy of deep learning models in reference [7] go against the original intention of lightweight distributed mining. Therefore, this article proposes a multi relationship network important node mining method based on passive clustering algorithm to solve the problems existing in existing methods. By using passive clustering algorithm, nodes are divided into different clusters according to their attributes, and attribute heterogeneity is considered during intra cluster analysis to avoid bias of a single global model; Constructing an adjacency matrix to dynamically describe network connections, combined with passive clustering mechanism to adaptively adapt topology changes, remove loosely related nodes, and maintain the compactness of the cluster structure; Adopting a clustering decomposition strategy to transform global problems into local intra cluster computations, reducing complexity and meeting the lightweight requirements of distributed mining; Combining the centrality of comprehensiveness and the characteristics of node attributes, evaluate the importance of nodes from both inter cluster and intra cluster perspectives, balancing the synergistic effects of structure and attributes. Experimental results have shown that this method can effectively improve the accuracy and efficiency of mining important nodes in multi relational networks, providing new theoretical support and practical paths for complex network analysis.

3 Design of important node mining method for multi-relational networks

3.1 Initial screening of important nodes in multi relational networks

Due to the existence of multiple types of relationships in the multi-relational network, and the relationship entities themselves have discrete characteristics, the binary relationship interaction structure between entities in the network is more complex, and it is difficult to intuitively track the interaction between entities corresponding to a variety of relationships [8, 9]. In this regard, the article constructs an adjacency matrix framework through the types of entities and relationships in the multi-relationship network, and represents the multi-relationship network as a matrix, so as to accurately describe the connectivity between relational entities (nodes) and clarify the corresponding resource allocation role of each node in the relational network.

Take an example of an unprivileged and undirected multi-relational network, which contains $\vec{n}$ relational entities (i.e., network nodes), a total of $\vec{m}$ relationship types, and a total of $\vec{c}$ network levels. The structure of this powerless and undirected multi-relational network is shown in Figure 1.

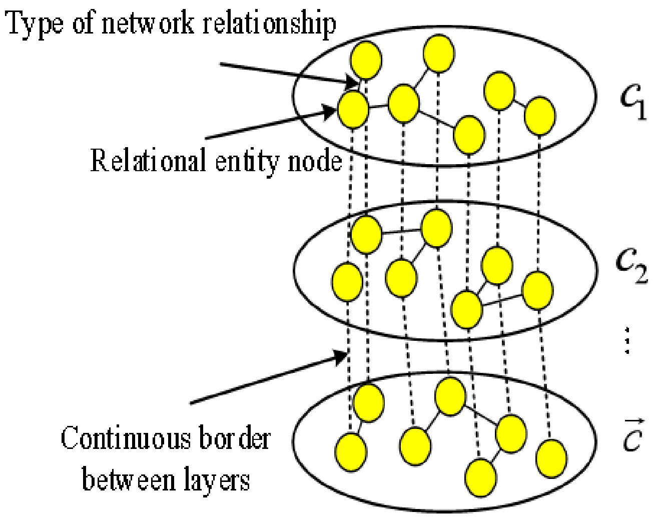

Figure 1: Unprivileged and undirected multi-relational network structure.

Based on Figure 1, it can be seen that each layer of the network structure in the powerless and directionless multi-relational network is homogeneous, and there are complex inter-layer connecting edge relationships between nodes of different network levels [10].

To make the network structure clearer, transform the multi relational network into a mathematical matrix. Among them, the intra layer adjacency matrix describes the connection relationship between nodes, as shown in the following equation:

$$\vec{Q}_1 = q_{l,i} = \sum_{g \in G_{\vec{m}}}^{\vec{n}} \varpi_g \times R'_g\left(l, i\right) \tag{1}$$

In Formula 1: $\vec{Q}_1$ denotes the intra-layer adjacency matrix of a multi-relational network; $q_{l,i}$ denotes the connection relationship between entity nodes l, i within the same layer; $G_{\vec{m}}$ denotes the set of relationship types within an unweighted and undirected multi-relational network; ϖ_g denotes the relative importance of relationship g in determining node connectivity, i.e., the weight coefficients corresponding to the relationship; $R'_g\left(l, i\right)$ denotes the function for calculating the undirected relationship of nodes within a layer.

The inter layer adjacency matrix describes the correlation between nodes across layers, as shown in the following equation:

$$\vec{Q}_2 = \vec{Q}_0 \oplus q_{l,i'}\left(c_1 + c_2\right) = \sum_{g \in G_{\vec{m}}}^{\vec{c}} \varpi_g \times R''_g\left(l, i'\right) \tag{2}$$

In Formula 2: $\vec{Q}_2$,$\vec{Q}_0$ denotes the inter-layer adjacency matrix of a multi-relational network with a weighted adjacency matrix; $\oplus$ denotes the weighting of inter-layer connection operator; $q_{l,i'}\left(c_1 + c_2\right)$ denotes the connection relationship between entity nodes l, i' in different levels c_1, c_2; $R''_g\left(l, i'\right)$ denotes the node undirected relationship calculation function in different levels.

The multi relational network adjacency matrix concatenates multiple layers of matrices into a block matrix, visually displaying the cross layer connection topology, as shown in the following equation:

$$\tilde{Q} = \left(\vec{Q}_1 + \vec{Q}_2\right)^{\vec{n}\vec{c} \times \vec{n}\vec{c}} = \begin{bmatrix} q_{11}^{1} & \cdots & q_{1n}^{\vec{m}} \\ \vdots & \ddots & \vdots \\ q_{\vec{n}1}^{\vec{m}} & \cdots & q_{\vec{n}\vec{m}}^{\vec{m}} \end{bmatrix} \begin{bmatrix} c_1 \\ \\ c_2 \\ \vdots \\ c_{\vec{c}} \end{bmatrix} = \left[c_1 q_{11}^{1} + \cdots + c_{\vec{c}} q_{\vec{n}\vec{m}}^{\vec{m}}\right] \tag{3}$$

In Formula 3: $\vec{Q}$ denotes the multi-relational network adjacency matrix; $q_{\vec{n}\vec{n}}^{\vec{m}}$ denotes the connection relationship between specific entity nodes with respect to the $\vec{m}$ th relationship type; $C_{\vec{c}}$ denotes the $\vec{c}$

th multi-relational network level; $C_{\vec{c}}q_{\vec{n}\vec{n}}^{\vec{m}}$ denotes the connection relationship between nodes in a specific network level with respect to a specific relationship type.The above formula uniformly describes the complex network structure in matrix form, quantifies the connection strength of different relationship types, and transforms the complexity of multi relationship networks into computable mathematical objects, providing quantitative basis for screening.

However, the matrix form only provides a quantitative description of the network structure [11, 12, 13]. To further explore important nodes, two key issues need to be addressed: the contribution of different relationship types to node importance varies, and bias caused by subjective weight setting needs to be avoided; Directly filtering nodes based on matrix elements may retain a large number of low activity nodes, increasing the redundancy of subsequent analysis. To this end, the entropy weight method is introduced to determine the weights of relationship types, and combined with the centrality of comprehensiveness and dynamic threshold screening, an effective mapping from matrix structure to key nodes is achieved. The entropy weight method can objectively quantify the importance of various relationships and avoid subjectively setting weights.The specific steps are as follows:

(1) Comprehensive centrality calculation

Weighted summation $\Gamma(l) = \sum_{i=1}^{\vec{n}} q_{l,i}$ is performed on the connection relationships between entity nodes l, i within the hierarchy to obtain the centrality of comprehensiveness $\Gamma(l)$. However, based on the connection relationships between nodes of the multi-relationship network presented by the elements in $\vec{Q}$, it is known that a single entity node may have connections to multiple types of relationships[14, 15, 16]. And node degree centrality simply takes into account the number of connections of nodes without distinguishing what type of relationship these connections are based on.Therefore, in the process of calculating node degree centrality, the article introduces the relationship type weights contained in $c_{\vec{c}}q_{\vec{n}\vec{n}}^{\vec{m}}$ to calculate the comprehensive degree centrality of nodes under multiple relationship types:

$$\Gamma'(l) = \sum_{g=1}^{\vec{m}} \varpi_g \sum_{i=1}^{\vec{n}} c_1 q_{li}^{g} \tag{4}$$

In Formula 4: $\Gamma'(l)$ denotes the centrality of the degree of integration of the entity nodes of the multi-relational network.

(2) Dynamic threshold filtering

After traversing and calculating all nodes within $\vec{Q}$ using Equation (5), due to the large number of nodes in the unweighted undirected multi relationship network and the wide distribution of degree centrality values among each node, a screening threshold is set based on the statistical characteristics of the comprehensive degree centrality of all nodes to achieve the initial screening of important nodes in the network. The expression for the filtering threshold is:

$$\hat{\theta} = \left\| \gamma_1 + \gamma_2 \right\| \tag{5}$$

In Formula 5: γ_1, γ_2 denote the mean and standard deviation of the centrality of the composite degree of all nodes, $\gamma_1 = \frac{1}{\vec{n}} \sum_{l=1}^{\vec{n}} \Gamma'(l), \gamma_1 = \frac{1}{\vec{n}} \sum_{l=1}^{\vec{n}} \Gamma'(l)$; $\widehat{\theta}$ denotes the potentially important node screening threshold. Based on the threshold calculation results,Based on the threshold calculation results, nodes with degree centrality greater than $\widehat{\theta}$ are identified as potential important nodes. After traversing all nodes in the matrix, the potential important nodes are sorted according to degree centrality, and a candidate set $\eta = [a_1, a_2, ..., a_{\hat{n}}]$ of potential multi relationship network important nodes is output, where $a_1, a_2, ..., a_{\hat{n}}$ represents potential important nodes containing relationship type information and $\widehat{n}$ represents the number of potential important nodes in the candidate set. Finally, the initial screening of important nodes is achieved.

3.2 Important node mining of multi-relational network based on passive clustering algorithm

After completing the initial screening of potential important nodes based on entropy weight method and comprehensive centrality, this paper further introduces passive clustering algorithm to finely analyze the screening results. The passive clustering algorithm achieves efficient and accurate mining of

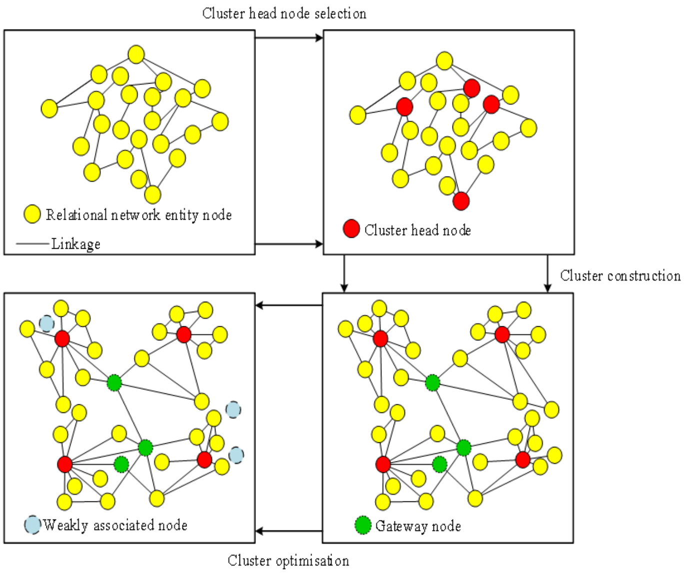

Figure 2: Passive clustering logic for multi-relational network nodes.

important nodes through dynamic adaptive clustering, local structure optimization, and multi relationship attribute fusion. Compared with traditional static clustering (such as K-means) and complex deep learning models, it is particularly suitable for dynamic multi relationship network environments. Firstly, based on the centrality of comprehensiveness, cluster head nodes are selected, and non cluster head nodes are dynamically clustered through correlation calculation. Loose nodes within the cluster are removed to optimize the local structure [17]; Subsequently, the importance of nodes is evaluated from both inter cluster (gateway node cross cluster connectivity) and intra cluster (weighted fusion of betweenness centrality, proximity centrality, and comprehensiveness centrality) perspectives. Finally, through threshold comparison, important nodes in a multi relationship network that combine local tightness and global influence are identified. This process not only inherits the global perspective of initial screening, but also deepens the ability to handle local heterogeneity through clustering mechanism. The specific process is as follows:

(1) Cluster head selection and dynamic clustering based on comprehensive centrality

Based on the distribution characteristics of cluster head nodes, the algorithm divides the selected important nodes into clusters according to their similarity characteristics [18, 19], facilitating unified resource allocation and task scheduling management for similar nodes within the cluster. And within each cluster, passive clustering algorithms follow local optimization logic. The passive clustering logic of multi relationship network nodes based on comprehensive centrality is shown in Figure 2:

As can be seen from Figure 2, the passive clustering of multi-relational network nodes mainly includes cluster head node selection, cluster construction, and the cluster optimisation process, and these clusters represent aggregates of multi-relational network entity nodes based on specific relationships [20]. In cluster head node selection, based on the multi-relationship network scale, the top U entity nodes with higher centrality of composite degree are selected in η, and these nodes will be used as the starting point of clustering, i.e., cluster head nodes. The result of cluster head node selection is expressed as:

$$\iota = \left[a_1', a_1', ..., a_U'\right] \tag{6}$$

$$U = \sqrt[m]{(\vec{n}+\vec{c})^2\,(\vec{n}-\vec{c})^2} \tag{7}$$

In Formula 6-7: ι denotes the set of selected cluster head nodes; $a_1', a_1', ..., a_U'$ denote the cluster head nodes.

Constructing clusters based on ι, passive clusters identify potentially multirelational network important nodes. In this process, in order to optimise the intra-cluster structure, for the non-cluster

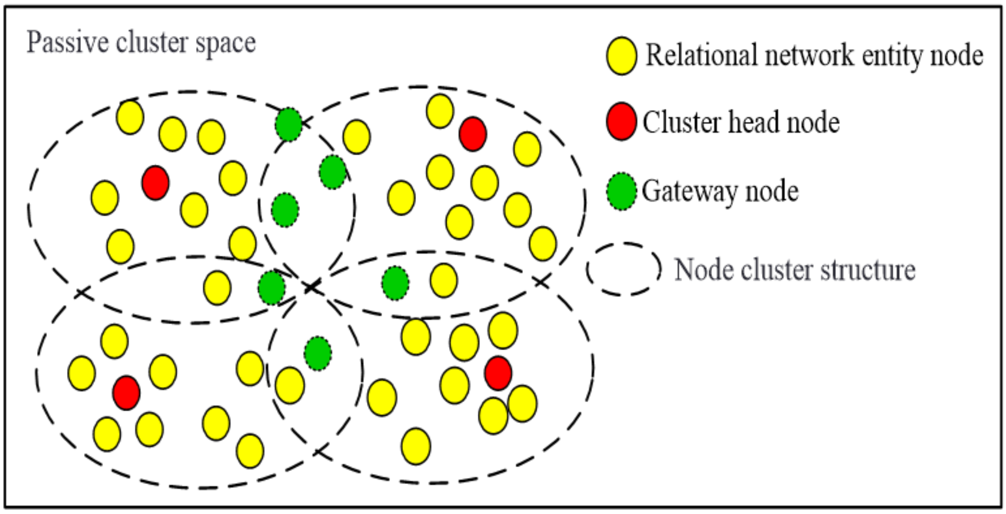

Figure 3: Schematic diagram of distribution state of gateway nodes after passive clustering.

head nodes in η, it is necessary to calculate the degree of association between them and the cluster head nodes, to measure the degree of dependence of the non-cluster head nodes on the cluster head nodes, and at the same time, assess the importance of the non-cluster head nodes within the cluster, and discover some potential key nodes that have a low degree of centrality, but have a high degree of association with the cluster head nodes. The expression for association degree calculation is as follows:

$$\hat{L}\left(a_o, a'_{o'}\right) = \sum_{g \in G_{\bar{m}}} \varpi_g \times R'_g\left(a_o, a'_{o'}\right) \tag{8}$$

In Formula 8: $\hat{L}\left(a_o, a'_{o'}\right)$ denotes the cluster node association degree; $a_o, a'_{o'}$ denotes the non-cluster head node with the cluster head node at a particular ordinal position.

Secondly, to ensure the tightness of the cluster structure, based on the multi-relational network scale and expert knowledge, the association threshold is defined, and the association between all non-cluster head nodes and the cluster head nodes is traversed and calculated and the results are compared with the set threshold, if the result of the calculation is greater than the set threshold, the non-cluster head node will be added to the corresponding cluster

$$B_{o'}$$

of cluster head node $a'_{o'}$ [21, 22]. The process is repeated continuously to expand the clusters until no node satisfies the adding condition, and all clusters $B_1, B_2, ..., B_U$ are obtained.

Finally, the sum of the correlation degree between all the nodes in the cluster two by two is averaged, and the average result is compared with the correlation degree of the nodes in the cluster, if the average correlation degree between the node and other nodes is low, the node needs to be removed from the cluster, so as to ensure that there is enough correlation degree between the nodes in the cluster, to avoid the existence of nodes whose relationship is too loose within the cluster, and to achieve the optimisation of the nodes of the multirelational network in the passive clustering. The distribution state of network nodes after passive clustering is shown in Figure 3.

From Figure 3, it can be seen that after the multi-relational network clustering, the gateway nodes are relatively dispersed in the edge position of each cluster, connecting different cluster nodes and establishing connections with multiple clusters at the same time, with multi-relational connection characteristics.

(2) Importance assessment of dual view nodes between and within clusters

Quantify node importance from both global (inter cluster) and local (intra cluster) dimensions, and discover key nodes that possess both local activity and global influence [23]. In the evaluation of inter cluster node importance, the first step is to traverse all clusters and use nodes connecting different clusters as gateway nodes, which can aggregate information related to relationships within different clusters [24, 25]. The gateway node is used as the subject of inter cluster importance analysis, and the calculation process of its inter cluster importance score is as follows:

$$\vec{K}\left(B_1, B_2\right) = \kappa_0 \sum_{\nu,\nu'=1}^{\kappa_0} \left(\vec{S}_\nu + \vec{S}_{\nu'}\right)^{1+\psi} \tag{9}$$

$$\hat{E}\left(a_{\overline{m}}\right) = \varphi \times \vec{K}\left(B_1, B_2\right) \tag{10}$$

In Formula 9-10: $\vec{K}(B_1, B_2)$ denotes the cross-cluster connectivity metric, i.e., the quantitative value of the range of influence of cross-cluster nodes (gateway nodes) under multi-relational network connectivity; κ_0 denotes the number of clusters connected by gateway nodes; $\vec{S}_v, \vec{S}_{v'}$ denotes the size of the node connecting to the ν, ν' th cluster; ψ denotes the dynamic evolutionary time scale of the multi-relational network; $\widehat{E}(a_{\bar{m}})$ denotes the inter-cluster importance score of the $\bar{m}-$th gateway node (belonging to the part of the nodes that are not cluster head nodes); φ denotes the proportion of the average size of the clusters, which belongs to the adjustment coefficients in Equation (10).

In the intra-cluster node importance assessment, the integrated degree centrality, median centrality and proximity centrality of the cluster nodes are used as the indicators for the intra-cluster node importance assessment. Since the composite degree centrality of the constructed cluster nodes has been calculated in the node traversal based on formula (5), it will not be repeated in the process of calculating the intra-cluster importance score. Taking cluster v as an example, a_o as the node to be assessed and a_1, a_2 as an arbitrarily selected intra-cluster node, the importance score of the node to be assessed within this cluster is calculated:

$$\delta\left(a_o\right) = \sum_{a_1 \neq a_2 \neq a_o \in \eta} \bar{s}_{a_0}\left(a_1, a_2\right) \times \bar{s}\left(a_1, a_2\right)^{-1} \tag{11}$$

$$\overline{\delta}\left(a_o\right) = \frac{1}{\sum_{a_1 \in \eta} \vec{d}\left(a_o, a_1\right)} \tag{12}$$

$$\tilde{E}\left(a_o\right) = \omega_1' \delta\left(a_o\right) + \omega_2' \overline{\delta}\left(a_o\right) + \omega_3' \Gamma'\left(a_o\right) \tag{13}$$

In Formula 11-13: $\delta(a_o)$ denotes the node median centrality within the cluster, i.e., the frequency of the node to be evaluated being located on the shortest path between other pairs of nodes; $\bar{s}_{a_0}(a_1, a_2)$ denotes the number of shortest paths from node a_1 to node a_2 and passing through the node to be evaluated; $\bar{s}(a_1, a_2)$ denotes the number of shortest paths between nodes; $\bar{\delta}(a_o)$ denotes the node proximity centrality, i.e., the proximity of the node to be evaluated to all other nodes; $\vec{d}(a_o, a_1)$ denotes the shortest distance between nodes; $\tilde{E}(a_o)$ denotes the importance score of the node to be evaluated within the cluster; $\omega_1^{'}, \omega_2^{'}, \omega_3^{'}$ denotes the weight coefficients corresponding to median centrality, proximity centrality, and composite degree centrality.

Linking formula (10) with formula (13), the combined node importance score $e(a_o)$ is calculated as follows:

$$e\left(a_o\right) = \tilde{E}\left(a_o\right) + \hat{E}\left(a_{\overline{m}}\right) \tag{14}$$

Based on formula (14) traversing all cluster nodes, and at the same time setting an absolute comprehensive importance score threshold tsy based on the a priori knowledge of multi-relational networks, the calculated comprehensive importance score of nodes is compared with the threshold, and important nodes of multi-relational networks with scores exceeding the threshold limit are mined.

4 Experimental analysis

4.1 Experimental environment construction

In order to verify the practical application effect of the designed multi-relational network important node mining method based on the passive clustering algorithm, the real multi-relational network dataset of CKM Physicians Innovation is selected, and a simulation test is carried out under the

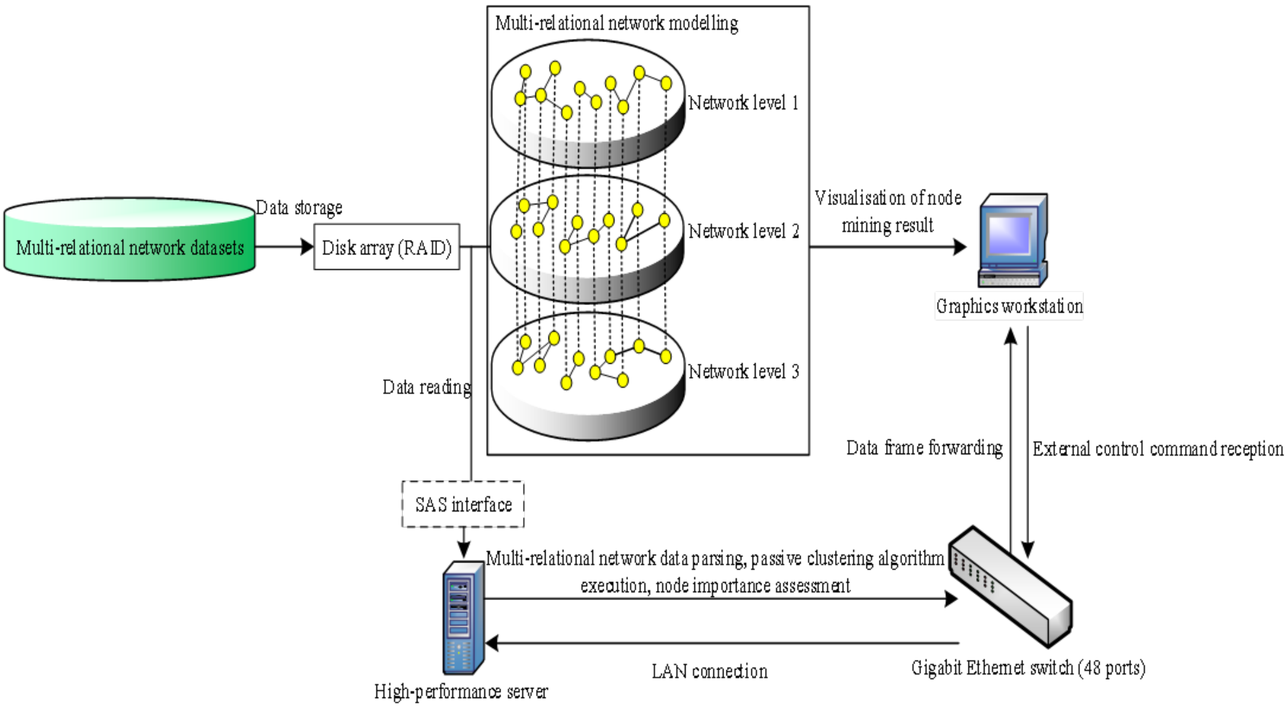

Figure 4: Layout of simulation experiment environment.

Table 1: Simulation test parameters

Parameter category	Simulation parameter name	Parameter value/description
Multi-relational network topology parameters	Network type	Social network
	Number of network layers	3 layers
	Total number of entity nodes	246
	First order distance of node degree	18.6
	Second order distance of node degree	405.4
	Number of network relationships	1551
	Type of node degree distribution	Poisson distribution
	Average clustering coefficient	0.42
Passive clustering algorithm parameters	Initial number of clusters	17
	Cluster merging threshold	≤0.46
	Intra-cluster similarity threshold	≥0.72
	Upper limit of inter-cluster connection weights	0.3
	Simulation time compensation	1s
	Number of iterations	50 times

environment of the Ubuntu 20.04 LTS system, and the codes are all implemented by CPython 3.8.10 in the process of the test, specifically, the layout of the simulation experiment environment is shown in Figure 4. Based on Figure 4, the simulation environment is built, while referring to the test running environment, describing some of the simulation parameters as shown in Table 1.

This experiment is based on social networks, simulating complex interactive relationships such as cooperation, consultation, and information dissemination between doctors in real medical scenarios. This experiment sets up three layers of relationships, corresponding to three interactive modes of cooperation, consultation, and information dissemination among doctors. This experimental network consists of 246 entity nodes, representing doctors or patients involved in diagnosis and treatment. In this experiment, the average degree is 18.6 and the second-order distance is 405.4, indicating that the node degree distribution is concentrated and there are some highly active nodes. This experiment contains 1551 relationship edges, reflecting the frequency of interaction between doctors. The experimental node degrees follow a Poisson distribution, indicating that the majority of node degrees are close to the average, while a few node degrees deviate significantly. The average clustering coefficient is 0.42, indicating significant community structure in the network. Based on network topology parameters and empirical rules, 246 nodes correspond to clusters 17 25, with the median value of 17 taken in this paper. Through pre experimental adjustments, the 0.46 threshold can avoid excessive merging and preserve local structural differences. A threshold of 0.72 ensures tight connections among nodes within the cluster and avoids interference from noisy nodes. Analyzing the distribution of inter cluster connections in the original dataset, 0.3 can filter out weak connections and preserve significant cross cluster interactions. Set 50 iterations to ensure the algorithm converges fully.

4.2 Experimental data

The dataset contains three layers of relationships between doctors, including cooperation, consultation, and information dissemination, covering 246 entity nodes and 1551 edges, truly reflecting the multidimensional interaction patterns of medical teams. Using the entity nodes and associated information in the experimental dataset as test samples, the training set and test set are randomly divided in an 8:2 ratio. Based on the experimental dataset, create datasets of different sizes including 10 nodes, 50 nodes, 90 nodes, 130 entity nodes, and relationship information.

$$w = \frac{w - \min\left(W_k\right)}{\max\left(W_k\right) - \min\left(W_k\right)} \tag{15}$$

In Formula 15: W_k denotes the weight set of the $k-th$ layer relationship.

4.3 Experimental indicators

(1)Accuracy of importance measurement

Measuring the degree of conformity between the algorithm's recognition of important node ranking and actual labeling is the core indicator for evaluating the effectiveness of node importance mining.

(2)Space complexity

Due to the fact that the data in the test dataset is represented as double precision floating-point numbers in programming languages (reasonable memory usage is 8 bytes), based on this data structure and memory usage relationship, the advantage range of node memory usage is defined as 80KB-214KB. By testing the memory usage of different methods when processing datasets of different sizes, the spatial complexity of the methods is analyzed, and the scalability of the methods is verified.

(3)Node importance calculation time

In multi-relational network important node mining, node degree centrality $\Gamma(l)$, median centrality $\delta(a_o)$ and proximity centrality $\bar{\delta}(a_o)$ are the basic indexes to measure the importance of a node, and at the same time, the calculation of their indexes is the main task in the node importance operation.

Test the running time ratio of calculating node degree centrality, betweenness centrality, and closeness centrality using different methods at different iteration times (the proportion of running time consumed by a method executing a single computational task to the total running time).

4.4 Experimental results

(1)Accuracy testing of importance measurement for multiple relational network nodes

In order to verify the effectiveness of the design method in mining important nodes, the accuracy of node importance measurement was tested. The methods in reference [5] and [6] were introduced as comparative methods to calculate the importance of the test set nodes. After sorting the calculated importance and corresponding nodes, the top 8 entity nodes with the highest importance ranking were extracted and verified against the actual importance ranking nodes. The specific test results are shown in Table 2:

The design method only has a unique error at the fourth node, which is due to the reverse sorting of nodes 81 and 82 due to their close correlation within the local cluster. However, reference [5] mistakenly placed low degree nodes 75 and 143 in the top 8 because they only rely on single-layer

Table 2: Multi-relational network node importance metric accuracy test results

Significant node sorting results (positive order)	Way of measuring			
	Actual result	Design method	Literature[5] method	Literature[6] method
1	Node 124	Node 124	Node 75	Node 30
2	Node 126	Node 126	Node 126	Node 121
3	Node 158	Node 158	Node 143	Node 94
4	Node 82	Node 81	Node 82	Node 82
5	Node 105	Node 105	Node 105	Node 201
6	Node 79	Node 79	Node 81	Node 79
7	Node 36	Node 36	Node 36	Node 215
8	Node 91	Node 91	Node 92	Node 91

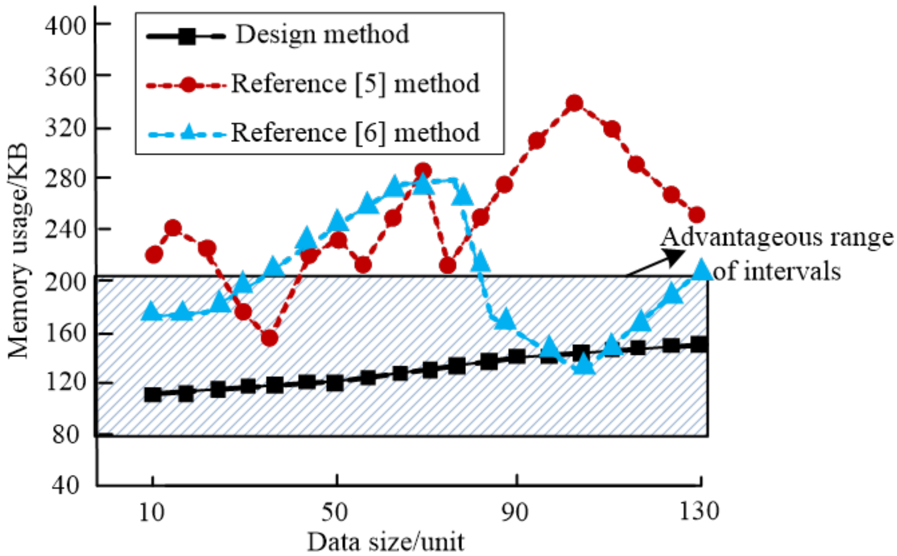

Figure 5: Space Complexity Test Results of Important Node Mining.

relationship centrality and ignore multi relationship fusion. Reference [6] misjudged nodes 30 and 94 as key, as the model overfits the local structure and ignores the "weakly connected but highly influential" bridging nodes in the medical network. From this, it can be seen that using design methods can effectively handle complex dynamic relationships in unweighted undirected networks, combine relationship weights to accurately measure node importance, effectively mine entity nodes with high importance in the network, and have good practical application effects.

(2) Space complexity testing of important node mining

In order to verify the scalability of the design method, the methods in reference [5] and reference [6] were introduced as comparative methods. The specific test results are shown in Figure 5:

As shown in Figure 5, with the continuous increase in data processing scale, the memory usage generated by each method also shows an overall continuous increase. However, when using design methods to increase the data size, the memory usage required for the method to run remains within a range with operational advantages, overall below 160KB, and the numerical results are superior to other methods. The memory growth in the method of reference [5] shows a fluctuating linear trend, while the memory change in the method of reference [6] shows a non-linear fluctuation. After 90 nodes, the memory unexpectedly drops to 130KB, but the overall standard deviation is as high as 47.2 KB. This is because the static clustering in reference [5] requires pre allocation of a fixed size adjacency matrix, resulting in memory occupancy proportional to the square of the number of nodes. The algorithm in reference [6] requires temporary storage of intermediate results during cluster merging, resulting in a significant increase in memory usage. From this, it can be seen that when dealing with the task of mining important nodes in multi relational networks, as the network size increases,

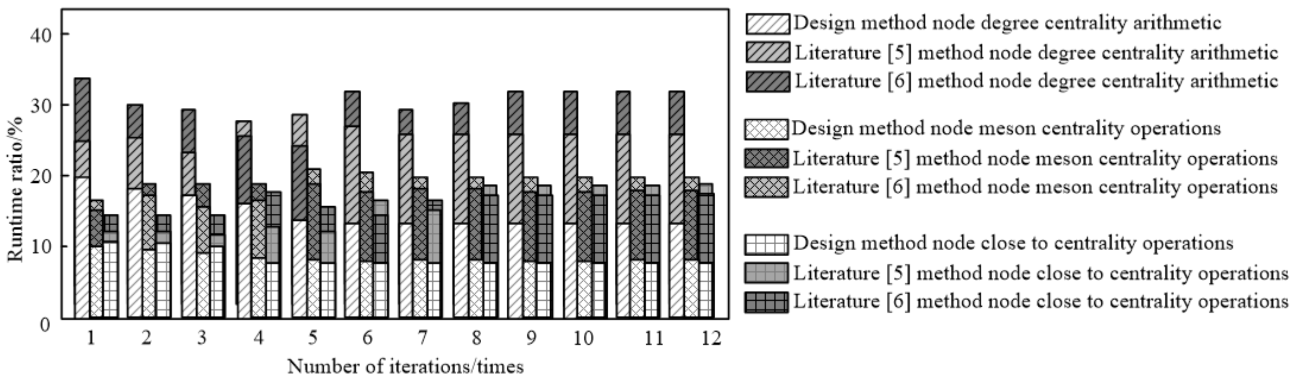

Figure 6: Nodal Importance Operation Time Test Results.

the memory required by the algorithm is still relatively small, exhibiting lower spatial complexity. Not only does it have good scalability, but it can also effectively cope with the dynamic evolution of relationship types and entities in multi relationship networks with its efficient node mining efficiency. This is because the proposed method filters low correlation nodes in advance through the intra cluster similarity threshold, reducing the subsequent computational workload.

(3)Node importance calculation time test

In order to test the efficiency and performance of the design method, the methods in reference [5] and reference [6] were introduced as comparison methods, and the importance calculation index values of teammate nodes in the iterative convergence state of each method were statistically compared. The specific test results are shown in Figure 6:

From Figure 6, it can be seen that during the entire process of node importance calculation using various methods, the computation time for node degree centrality is relatively long, that is, the computation time is mainly consumed in degree centrality calculation. The difference in computation time between node betweenness centrality and near centrality is small. The running time ratio for node degree centrality operation using the design method is 14.2%, for betweenness centrality operation is 8.9%, and for near centrality operation is 8.6%. The numerical results are all superior to other methods, indicating that the method has good efficiency performance. This is because the design method decomposes the global network into multiple local clusters by dynamically adjusting the cluster structure, and independently calculates the importance of nodes within each cluster, significantly reducing computational complexity. The running time ratio of the method in reference [5] for node degree centrality operation is 26.5%, for betweenness centrality operation is 15.3%, and for near centrality operation is 12.1%. It is because this method requires recalculating the cluster center coordinates every time the clustering is adjusted, resulting in a waste of time. The running time ratio of the method in reference [6] for node degree centrality operation is 32.8%, for betweenness centrality operation is 19.7%, and for near centrality operation is 14.5%. It is because this method requires the introduction of a Dropout layer to avoid overfitting, but the pruning operation also requires readjustment of the network structure, which increases the time.

5 Result

This article comprehensively evaluates the performance advantages of the proposed method for mining important nodes in multi relational networks based on passive clustering algorithm through systematic experimental verification.

(1) In the accuracy test of node importance measurement, the recognition accuracy of the first 8 nodes in the design method is only 1 rank error, which is significantly better than literature [5] and literature [6]. This indicates that the comparative method leads to misjudgment of key nodes due to neglecting multi relationship fusion and overfitting local structures, while the design method effectively solves attribute heterogeneity interference through entropy weight fusion and clustering evaluation.

(2) The space complexity test shows that when the network scale expands to 130 nodes, the design method maintains a stable memory occupancy of 150KB. Its fundamental advantage lies in

the dynamic filtering of the passive clustering structure decoupling entropy weight method, while the comparative method has a memory occupancy of over 300KB due to static matrix pre allocation and intermediate result cache expansion.

(3) The centrality of the design method is 14.2%, which is 18.6% lower than that in reference [6], the betweenness centrality is 8.9%, which is 6.4% lower than that in reference [5], and the near centrality is 8.6%, which is 3.5% lower than that in reference [5]. This is due to the dynamic clustering optimization of independent computing mechanisms within the cluster.

Experimental results have shown that this method outperforms existing methods in terms of accuracy, memory efficiency, and computational speed by decoupling the structure of passive clustering, filtering relationships using entropy weight method, and accurately measuring hierarchical evaluation. It provides an efficient and reliable solution for key node mining in multi relationship dynamic networks.

6 Conclusions

The article proposes a method for mining important nodes in multi relational networks based on passive clustering algorithm. By constructing an adjacency matrix framework to quantify the network structure, combined with comprehensive centrality to preliminarily screen candidate nodes, and innovatively introducing passive clustering algorithm into multi relationship network analysis, heterogeneity processing of node attributes and local structure optimization are achieved. The experimental results show that this method has a stable memory usage of less than 160KB when dealing with undirected networks without authority, and the computation time ratios of node degree centrality, betweenness centrality, and proximity centrality are reduced to 14.2%, 8.9%, and 8.6%, respectively, which is significantly better than traditional methods. This method can be applied in different fields such as healthcare, finance, and social networks, and has excellent adaptability. However, sparse matrix operations have not been combined with dedicated hardware acceleration such as FPGA, and there is still room for optimization in degree centrality computation time. In future research, we will focus on dynamic attribute weight learning, algorithm hardware collaborative acceleration, and cross modal network extension to further promote the practical application and performance improvement of complex network analysis technology.

Funding

The APC was funded by R&D center "Cercetare Dezvoltare Agora" of Agora University.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Y. Yue, G. Wang, J. Hu, and Y. Li, “An improved label propagation algorithm based on community core node and label importance for community detection in sparse network,” *Applied Intelligence*, vol. 53, no. 14, pp. 17 935–17 951, 2023.

[2] L. Zhang, S. Wang, J. Liu, X. Chang, Q. Lin, Y. Wu, and Q. Zheng, “Mul-grn: Multi-level graph relation network for few-shot node classification,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 6, pp. 6085–6098, 2022.

[3] Y. Guo, "Towards the efficient generation of variant design in product development networks: network nodes importance based product configuration evaluation approach," *Journal of Intelligent Manufacturing*, vol. 34, no. 2, pp. 615–631, 2023.

[4] Y. Jiang, "A study of trust mining algorithms for beacon nodes in large-scale network environments," *PeerJ Computer Science*, vol. 11, p. e2755, 2025.

[5] C. Song, X. Zhang, X. Liu, D. Ren, and M. Dong, "Key node mining algorithm for directed weighted air quality network based on propagation characteristics," in *Journal of Physics: Conference Series*, vol. 1693, no. 1. IOP Publishing, 2020, p. 012066.

[6] T. H. Zhou, B. Jiang, Y. Lu, and L. Wang, "An adaptive seed node mining algorithm based on graph clustering to maximize the influence of social networks," in *International Conference on Intelligent Computing.* Springer, 2020, pp. 498–509.

[7] B. Samya, B. Anantharam, L. Swathi, D. K. Sreeramamurthy, and M. V. Rao, "An intelligent distributed data mining framework for energy-efficient wsn using a hybrid heuristic-aided cascaded residual lstm," *Peer-to-Peer Networking and Applications*, vol. 18, no. 5, p. 252, 2025.

[8] J. Chen, W. Wang, K. Yu, X. Hu, M. Cai, and M. Guizani, "Node connection strength matrix-based graph convolution network for traffic flow prediction," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 9, pp. 12 063–12 074, 2023.

[9] W. Fan, F. Xiao, M. Lv, L. Han, J. Wang, and X. He, "Node essentiality assessment and distributed collaborative virtual network embedding in datacenters," *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 4, pp. 1265–1280, 2023.

[10] B. Niu, S. Zhou, and H. Zhang, "Influential node identification of network based on agglomeration operation," *International Journal of Foundations of Computer Science*, vol. 35, no. 03, pp. 271–295, 2024.

[11] J. R. Raja, J. G. Lee, D. Dhotre, P. Mane, O. S. Rajankar, A. Kalampakas, N. D. Jambhekar, and D. Bhalke, "Fuzzy graphs and their applications in finding the best route, dominant node and influence index in a network under the hesitant bipolar-valued fuzzy environment," *Complex & Intelligent Systems*, vol. 10, no. 4, pp. 5195–5211, 2024.

[12] X. Wen, B. Si, M. Xu, F. Zhao, and R. Jiang, "A passenger flow spatial–temporal distribution model for a passenger transit hub considering node queuing," *Transportation Research Part C: Emerging Technologies*, vol. 163, p. 104640, 2024.

[13] T. T. Huynh, C. T. Duong, T. T. Nguyen, V. T. Van, A. Sattar, H. Yin, and Q. V. H. Nguyen, "Network alignment with holistic embeddings," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1881–1894, 2021.

[14] R. Taub and Y. Savir, "Saf: Smart aggregation framework for revealing atoms importance rank and improving prediction rates in drug discovery," *Journal of Chemical Information and Modeling*, vol. 64, no. 10, pp. 4021–4030, 2024.

[15] F. Bian, A. G. Yeh, and J. Zhang, "Percolating spatial scale effects on the landscape connectivity of urban greenspace network in beijing, china," *Landscape and Ecological Engineering*, vol. 20, no. 1, pp. 33–51, 2024.

[16] G. J. Goodwin, C. E. Salva, J. Rodrigues, J. Maietta, H. C. Kuwabara, S. Ross, T. F. Kinsora, and D. N. Allen, "Characterizing the network structure of post-concussion symptoms," *Archives of clinical neuropsychology*, vol. 38, no. 5, pp. 690–698, 2023.

[17] Z. A. M. Shamki and F. Rabee, "Image mining technique using hadoop map reduce over distributed multi-node computers connections," *Al-Salam Journal for Engineering and Technology*, vol. 1, no. 2, pp. 18–24, 2022.

[18] S. D. Mishra and D. Verma, "Energy-efficient and reliable clustering with optimized scheduling and routing for wireless sensor networks," *Multimedia Tools and Applications*, vol. 83, no. 26, pp. 68 107–68 133, 2024.

[19] S. K. Gupta and D. P. Singh, "Seed community identification framework for community detection over social media," *Arabian Journal for Science and Engineering*, vol. 48, no. 2, pp. 1829–1843, 2023.

[20] C. Gupta, L. Singh, and R. Tiwari, "Malicious node detection in vehicular ad-hoc network (vanet) using enhanced beacon trust management with clustering protocol (ebtm-cp)," *Wireless Personal Communications*, vol. 130, no. 1, pp. 321–346, 2023.

[21] Z. Shou, M. Tang, H. Wen, J. Liu, J. Mo, and H. Zhang, "Key student nodes mining in the in-class social network based on combined weighted gra-topsis method," *International Journal of Information and Communication Technology Education (IJICTE)*, vol. 19, no. 1, pp. 1–19, 2023.

[22] S. Tripathi, O. J. Pandey, and R. M. Hegde, "Socially aware network clustering for throughput maximization in mobile wireless sensor networks," *IEEE Transactions on Network and Service Management*, vol. 21, no. 1, pp. 838–850, 2023.

[23] W. Zhang, J. Wang, G. Han, Y. Feng, and X. Tan, "A nonuniform clustering routing algorithm based on a virtual gravitational potential field in underwater acoustic sensor network," *IEEE Internet of Things Journal*, vol. 10, no. 15, pp. 13 814–13 825, 2023.

[24] G. Mahalakshmi, S. Ramalingam, and A. Manikandan, "An energy efficient data fault prediction based clustering and routing protocol using hybrid asso with mernn in wireless sensor network," *Telecommunication Systems*, vol. 86, no. 1, pp. 61–82, 2024.

[25] T. Panse and P. Panse, "An efficient gateway node selection method for clustering in heterogeneous mobile ad-hoc networks," *Wireless Personal Communications*, vol. 135, no. 1, pp. 113–126, 2024.

Cite this paper as:

Huang, H.J. (2026). A Passive Clustering-Based Approach for Important Node Mining in Multi-Relational Networks, *International Journal of Computers Communications & Control*, 21(2), 7108, 2026.
https://doi.org/10.15837/ijccc.2026.2.7108