Chapter 5

MNERLP-MUL: Merged Node and Edge Relevance based Link Prediction in Multiplex Networks

In the previous chapter, an attempt was made to perform link prediction on multiplex networks using influence spread across multi-hop paths between nodes. Though the 3-degree of influence phenomenon leads to taking into account paths as long as six hops, the node's role in the entire graph structure is not taken into account. In order to improve upon this issue, in this chapter, we attempt to combine node and edge relevance to enhance link prediction in multiplex networks and propose a method called Merged Node and Edge Relevance based Link Prediction in Multiplex Networks ($MNERLP - MUL$). This method follows the quasi-local information-based link prediction template, which attempts to find a trade-off between local (edge relevance) and global (node relevance) information. In this chapter[1], it is theorized that to accurately perform this link prediction, the relevance of both the edges as well as the nodes has to be taken into account.

[1]Published in Journal of Computational Science, DOI:https://doi.org/10.1016/j.jocs.2022.101606

5.1 Introduction

This chapter proposes a link prediction method on multiplex networks based on the combined relevance of nodes and edges present on indirect paths between nodes (common neighbor-based or 1-hop paths). The primary motivation behind the $MNERLP-MUL$ method is to use both local (edge relevance) and global information (node relevance) to perform link prediction. This is the principle behind quasi-local information-based link prediction methods in simple graphs, which attempt to combine the best characteristics of both local and global methods to perform more accurate link prediction. A density-based aggregation model is used to generate a summarized weighted graph from all the layers of the multiplex network. The inspiration for the $MNERLP-MUL$ method is based on the fact that quasi-local similarity methods, which make use of regional information at a larger scale than ones based on immediate local regions, are shown to be more accurate in the case of simple single-layered networks. Based on the 3 Degree of Influence Phenomenon (Christakis and Fowler[28, 29]), we believe that the overall relevance of an edge is primarily a function of local information such that nodes only influence it in its 3-hop neighborhood. On the other hand, node relevance is modeled as dependent on global information, which is influenced by the contribution of the node on the shortest paths between all nodes (centralities). Finally, based on edge and node relevance, an adequate score is assigned to the possibility of a link between unconnected nodes.

5.1.1 Edge Relevance - Local 3 Degree of Influence Model

The literature states that the influence of a node is spread to 3 hops from its origin (Christakis and Fowler[28, 29]), hence different levels of edge influence of a node A are considered as Level-1(A), Level-2(A), Level-3(A), Level-4(A) and Level-5(A) as shown in Fig.5.1 (represented by Lev.1, Lev.2, Lev.3, Lev.4 and Lev.5 in figure respectively). Level-1(A) are the edges connecting A with its direct neighbors (A-B, A-C, A-D) while

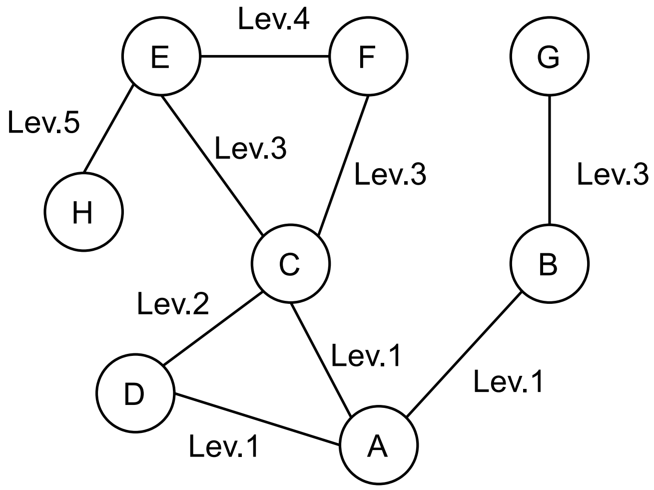


FIGURE 5.1: Edge level network structure of node A (Lev.1 is Level-1(A) and so on and so forth) demonstrating the influence node A exerts on edges 3 hop distances away from it.

Level-2(A) are edges between these direct neighbors (C-D). Level-3(A) are edges connecting direct neighbors with their indirect counterparts (B-G, C-E, C-F) while Level-4(A) are those between indirect neighbors (E-F). All the edges at a distance of 3 hops from A which do not belong to Level-4(A) are considered as the last circle Level-5(A) (E-H). The region of 3 hops from the node A is considered following the principle of three degrees of influence ([142]). Within this region, the product of importance and existing edge weight is used to quantize the influence of each particular level of ego network such that edges belonging to Level-1(A), Level-2(A), Level-3(A), Level-4(A) and Level-5(A) each have influence equal to $5*W, 4*W, 3*W, 2*W, 1*W$ respectively such that W is the weight of the edge under consideration.

5.1.2 Node Relevance - Global Centrality Model

To measure node relevance of each node, centrality measure is used. Three popular centrality measures have been used to compare and contrast the performance of $MNERLP-MUL$ approach which are calculated using different variations of shortest path distances between nodes. The centrality measure taken into consideration are as follows:

- *Closeness Centrality* - Closeness centrality (Freeman[163]) of a node x is the reciprocal of average shortest distance from a particular node to all other reachable nodes of the graph. Here $dist(x,a)$ is shortest distance between $x\&a$ and n is total number of nodes in graph $G(V,E)$.

$$CC(x) = \frac{n-1}{\sum_{a\in V\neq x} dist(x,a)} \tag{5.1}$$

- *Betweenness Centrality* - Betweenness centrality (Brandes[164]) of a node x is the sum of fraction of all pair shortest paths which pass through that particular node. Here $\gamma(a,b)$ is the number of shortest paths between $a\&b$ and $\gamma(a,b\|x)$ is the number of those shortest paths in which x exists.

$$BC(x) = \sum_{a,b\in V} \frac{\gamma(a,b \mid x)}{\gamma(a,b)} \tag{5.2}$$

- *Harmonic Centrality* - Harmonic centrality (Boldi and Vigna[165]) of a node x is the sum of reciprocal of shortest path distances from the node x to all other nodes of the graph.

$$HC(x) = \sum_{a\in V\neq x} \frac{1}{dist(a,x)} \tag{5.3}$$

5.2 Proposed Work

In this section, the proposed framework $MNERLP-MUL$ is discussed, which takes multi-interaction networks and uses merged node and edge based link prediction framework to predict missing links in these multiplex networks. The proposed framework is shown in Fig.5.2. The proposed algorithm, $MNERLP-MUL$, consists of three basic steps -

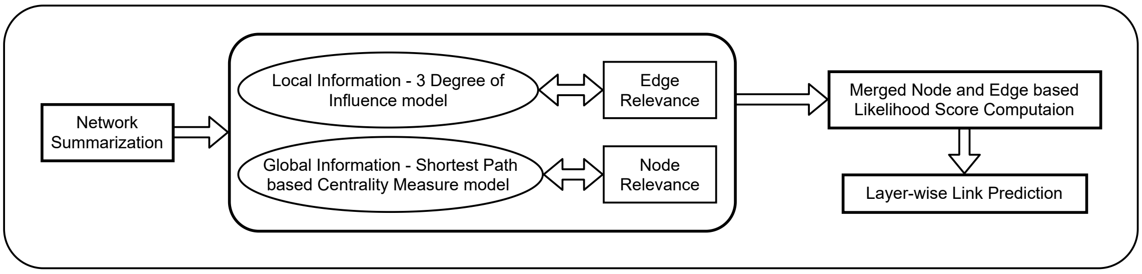


FIGURE 5.2: $MNERLP-MUL$ Framework demonstrating the overall structure of link prediction workflow.

- In the first step, we attempt to collate the dis-separate information from all layers into one summarized weighted graph. An aggregation model is modified to consider the differences in overall edge densities between layers as discussed in Section 5.2.1.
- Secondly, the edge relevance of existing edges is calculated (using 3 Degree of Influence theory as mentioned in Section 5.1.1) and node relevance (centrality measures as mentioned in Section 5.1.2) based on the current state of graph G_{MNERLP} as discussed in Section 5.2.2.
- The third step is to calculate the likelihood score of non existing edges using the edge and node relevance calculated in the previous step as discussed in Section 5.2.3.

5.2.1 Network Summarization

In $MNERLP-MUL$ approach, weights have been used to represent density dissimilarities between layers and weights also help in the nuanced transposition of edge probabilities from the summarized graph to actual layers. Network summarization is, in $MNERLP-MUL$ framework, the process of coupling multi interaction (multiplex) networks into a single weighted network. A topological integration approach has been used to form such a weighted network. Therefore, connection strength $A_M(x,y)$ of any existing edge (x,y)

in such network can be computed using Equation 5.4 [162]. Combining all edges which have some connection strength into a single graph, the graph G_M is produced.

$$A_M(x,y) \leftarrow \frac{1}{n}\sum_{j=1}^{n}\{A_j \mid A_j = [a^j_{xy}]_{|V|\times|V|}\} \tag{5.4}$$

where,

$$a^j_{xy} \leftarrow \begin{cases} 1 & \textit{if } \exists(x,y) \in E_j, j \in [1,n] \\ 0 & \textit{otherwise} \end{cases}$$

For the approach in $MNERLP - MUL$, a slightly modified method is proposed that considers relative densities of the layers amongst themselves. Its is theorized that using this in combination with a suitable calculation method at the time of re-transforming summarized graph probabilities to the original layers can achieve better results. The two introduced parameters for layer fusion (packing) and likelihood transposition (unpacking) are defined in Equations 5.5 and 5.6.

$$P(j) \leftarrow \frac{1}{\mid E_j \mid} \tag{5.5}$$

$$UP(j) \leftarrow \frac{\mid E_{G_{MNERLP}} \mid - \mid E_j \mid}{\mid E_{G_{MNERLP}} \mid} \tag{5.6}$$

The graph G_{MNERLP} will be the same as G_M, where nodes in these graphs have an edge if any of the layers have the same edge. The modified summarized weight matrix is as follows -

$$A_{MNERLP}(x,y) \leftarrow \frac{1}{n}\sum_{j=1}^{n}\left(\{A_j \mid A_j = [a^j_{xy}]_{|V|\times|V|}\} * P(j)\right) \tag{5.7}$$

5.2.2 Edge and Node Relevance Calculation

In Section 5.1.1 the discussion is on how the contributive effect of node influence spreads to the edges in its vicinity. Different levels of edge sets and quantified the effect of node on them have been shown, even when it is not directly edge adjacent (correlating nodes and edges using 3 Degree of Influence theory). When the cumulative effect of all nodes is taken into consideration, the total relevance of an edge (a,b) can be calculated with weight w using the following equation -

$$ER(a,b) = \sum_{n \in V} Level-Score(a,b,n) \tag{5.8}$$

such that,

$$Level-Score(a,b,n) = \begin{cases} 5*w, \mathit{iff}\ (a,b) \in Level-1(n) \\ 4*w, \mathit{iff}\ (a,b) \in Level-2(n) \\ 3*w, \mathit{iff}\ (a,b) \in Level-3(n) \\ 2*w, \mathit{iff}\ (a,b) \in Level-4(n) \\ 1*w, \mathit{iff}\ (a,b) \in Level-5(n) \\ 0, otherwise \end{cases} \tag{5.9}$$

For measuring the effect of node relevance a quantitative method has been used where node relevance is calculated using the centrality measure value of node in graph. Different shortest path-based centrality measures have been used (as shown in Section 5.1.2) and the node relevance for node x can be computed as -

$$NR(x) = \begin{cases} CC(x), \mathit{iff}\ Centrality \leftarrow Closeness \\ BC(x), \mathit{iff}\ Centrality \leftarrow Betweenness \\ HC(x), \mathit{iff}\ Centrality \leftarrow Harmonic \end{cases} \tag{5.10}$$

5.2.3 Likelihood score Computation

The likelihood of edge existence can be determined by collective effect of common neighbors on unconnected edges. This effect is divided in two separate parts i.e., node relevance based on the common neighbor themselves and edge relevance based on the effect of connecting edges between common neighbor and target nodes. Also two parameters are introduced which represent the relative weightage of node and edge relevance to the final link likelihood. The likelihood score $LI(u,v)$ of non-existing link u,v to predict missing link can be computed as follows using merged node and edge relevance.

$$LI(u,v) \leftarrow \sum_{z \in N(u) \cap N(v)} \left(\frac{MR_{CN}(u,z,v)}{\sum_{x \in N(z)} MR_{CNIN}(z,x)} \right) \tag{5.11}$$

where MR_{CN}&MR_{CNV} i.e., merged relevance based on single common neighbor and merged relevance based on common neighbor vicinity is,

$$\begin{aligned} MR_{CN}(a,b,c) \leftarrow & \left((ER(a,b) + ER(b,c))^{\alpha} * NR(b)^{\beta} \right), \\ & iff b \in N(a) \cap N(c) \end{aligned} \tag{5.12}$$

$$MR_{CNV}(d) \leftarrow \sum_{e \in N(d)} MR_{CNV}(d,e) = \sum_{e \in N(d)} \left(ER(d,e)^{\alpha} * NR(e)^{\beta} \right) \tag{5.13}$$

Finally this likelihood score or probability of an edge existing in the summarized weighted graph can be used to calculate probability of an edge in a particular layer by -

$$LI_j(a,b) \leftarrow LI(a,b) * UP(j) \tag{5.14}$$

Algorithm 3: MNERLP-MUL: Merged Node and Edge Relevance based Link Prediction in Multiplex Networks

Input: Social Networks: $G_i(V, E_i)$
Output: Likelihood Index: LI

1 ▷ Network Integration
2 Create a multiplex network A_{MNERLP} from n different interaction networks on same user set using Equation 5.7
3 ▷ Edge Relevance
4 Calculate Edge Relevance of existing edges using collective effect of 3 Degree of Influence of all nodes of graph ER by Equation 5.8
5 ▷ Node Relevance
6 Calculate Node Relevance of all nodes of graph NR by Equation 5.10
7 ▷ Merged Node and Edge Relevance based Computation Likelihood Index Computation
8 $LI_{|V|*|V|} \leftarrow 0$
9 **for** *each non-existing link* $(u,v) \notin E$ **do**
10 ▷ $LI[u][v]$ likelihood score of non-existing pair (u,v) calculated using Equation 5.11
11 **for** *each common neighbor* z *of* (u,v) **do**
12 Calculate $MR_{CN}(u,z,v)$ using Equation 5.12
13 $MR_{CNV}(z) \leftarrow 0$
14 **for** *each neighbor* x *of* z **do**
15 $MR_{CNV}(z) = MR_{CNV}(z) + MR_{CNV}(z,x))$ using Equation 5.13
16 $LI[u][v] = LI[u][v] + \frac{MR_{CN}(u,z,v)}{MR_{CNV}(z)}$
17 **Return** LI;

5.2.4 $MNERLP-MUL$ Algorithm with an illustrative example

The **Algorithm 3** demonstrates how the likelihood score matrix is calculated for graph G_{MNERLP}. The input to the algorithm is the layer graphs $G_i \mid i \in (1,n)$. The output is likelihood matrix of dimension $|V| * |V|$ such that $|V|$ is number of nodes in G_{MNERLP}. The algorithm can be divided into three major modules - initialization and graph summarization (lines 1-2), edge and node relevance calculation using summarized graph (lines 3-6) and lastly the calculation of likelihood score for non existent edges (lines 7-17). An empty matrix LI is defined in line 8 and each cell corresponds to a non-existent edge of graph G_{MNERLP}. For single non-existent edge, the effect of all common neighbors is calculated in lines 11-15 and $LI[u][v]$ is appropriately amended in

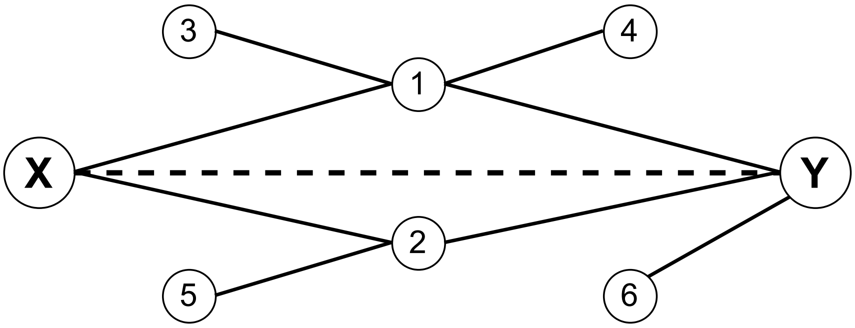


FIGURE 5.3: Example graph of $MNERLP-MUL$ based link prediction for calculation of $LI(X,Y)$

line 16. The final likelihood matrix must be normalized before calculation of final layer specific probabilities.

Fig 5.3 shows a demonstrative example of link prediction using $MNERLP-MUL$ between nodes $X\&Y$. It is assumed here that edge based relevance $ER(n_1,n_2)$ (which is a representation of edge importance using local information of nodes $n_1\&n_2$) and node relevance $NR(n)$ (which is representation of node importance calculated using global information for node n) is already calculated. For example, to calculate $ER(X,1)$, it can be seen (as relative to nodes of graph) that edge $X-1$ is part of $Level-1(X)$, $Level-1(1)$, $Level-3(3)$, $Level-3(4)$, $Level-5(5)$, $Level-5(6)$ and $Level-3(Y)$. All these relevant node influences (Equation 5.9) are combined to calculate $ER(X,1)$ (Equation 5.8). This notation form is explained in Section 5.1.1. $NR(n)$ is calculated using Equation 5.10 such that the whole graph structure (global information) is used for its calculation. Now for calculation of link likelihood between nodes $X\&Y$ we refer first to common neighbors of these nodes, i.e., nodes 1&2. The merged relevance of these nodes (line 7 and 10 of Algorithm 3) will be summed up to calculate final $LI(X,Y)$. The merged relevance of node 1 has two components - $MR_{CN}(X,1,Y)$ (using edges $X-1\&Y-1$ as well as the node itself) and $MR_{CNV}(1)$ the overall contribution of node 1 with respect to its neighbors (nodes $3,4,X,Y$). $MR_{CN}(X,1,Y)$ is calculated using Equation 5.12 and $MR_{CNV}(1)$ is calculated using Equation 5.13. $\frac{MR_{CN}(X,1,Y)}{MR_{CNV}(1)}$ is the current contribution of node 1 to $LI(X,Y)$ (line 11 of Algorithm 3). Similar procedure is followed for node 2. The total calculation of final $LI(X,Y)$ is given in Equation 5.15.

$$LI(X,Y) = \frac{MR_{CN}(X,1,Y)}{MR_{CNV}(1)} + \frac{MR_{CN}(X,2,Y)}{MR_{CNV}(2)} \tag{5.15}$$

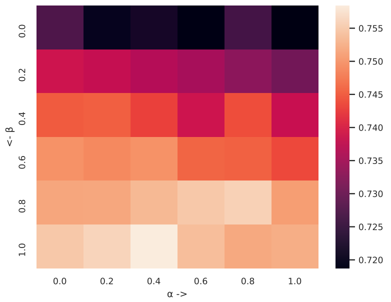
where,

$$\begin{aligned}
MR_{CN}(X,1,Y) &= \big(ER(X,1)+ER(1,Y)\big)^{\alpha} * NR(1)^{\beta} \\
MR_{CN}(X,2,Y) &= \big(ER(X,2)+ER(2,Y)\big)^{\alpha} * NR(2))^{\beta} \\
MR_{CNV}(1) &= \sum_{n_1 \in \{3,4,X,Y\}} \Big(ER(1,n_1)^{\alpha} * NR(n_1)^{\beta}\Big) \\
MR_{CNV}(2) &= \sum_{n_2 \in \{5,X,Y\}} \Big(ER(2,n_2)^{\alpha} * NR(n_2)^{\beta}\Big)
\end{aligned} \tag{5.16}$$

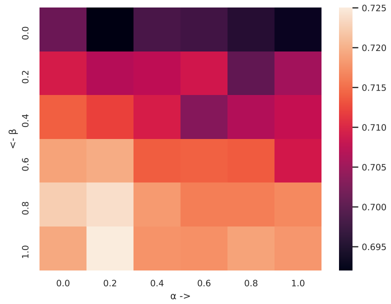
5.3 Performance Analysis

5.3.1 Parameter Variation Comparison

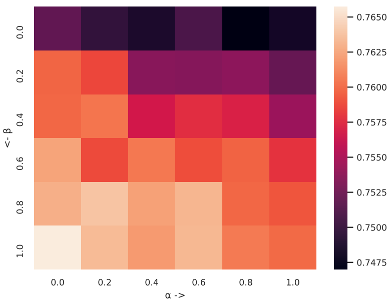
In this section, the comparison of the performance of the $MNERLP-MUL$ algorithm is shown for different values of $\alpha \& \beta$ parameters, which represent the relative weightage of edge and node relevance to the final likelihood calculation. In Fig.5.4, it is observed that in all datasets except CKM-Physicians-Innovation, the best AUC values are encountered when $\beta = 1.0$. For α, they correspond to a range of $0.0 - 0.4$. For the CKM-Physicians-Innovation dataset, though the best value of AUC is seen at $\beta = 0.8$ and α remains similar to the other three datasets at 0.2. A similar pattern is observed in Fig.5.5 and Fig.5.6 for balanced accuracy score and F1 score heatmaps. Hence, the best combination is $\alpha = 0.2$ and $\beta = 1.0$. In the experiments of this section, closeness centrality has been used as the measure for node relevance.



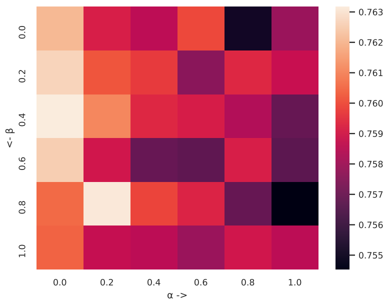
(A) Vickers-Chan-7thGraders



(B) Kapferer-Tailor-Shop

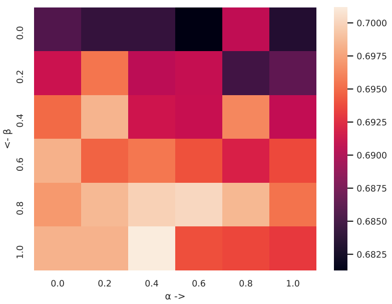


(C) Lazega-Law-Firm

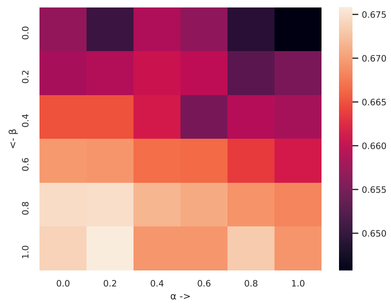


(D) CKM-Physicians-Innovation

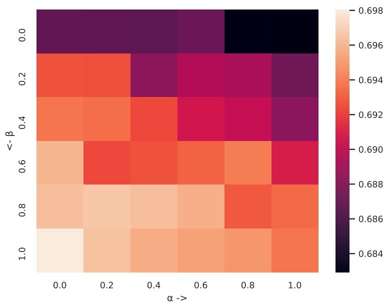
FIGURE 5.4: Heatmap of AUC variation of $MNERLP-MUL$ algorithm's performance with respect to α & β with *Ratio* of testing to total edges averaged over range $0.1-0.5$



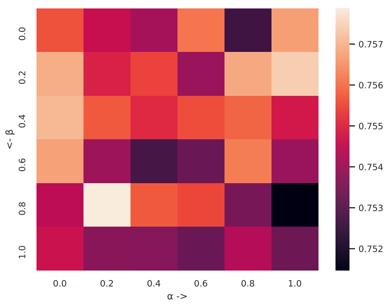
(A) Vickers-Chan-7thGraders



(B) Kapferer-Tailor-Shop

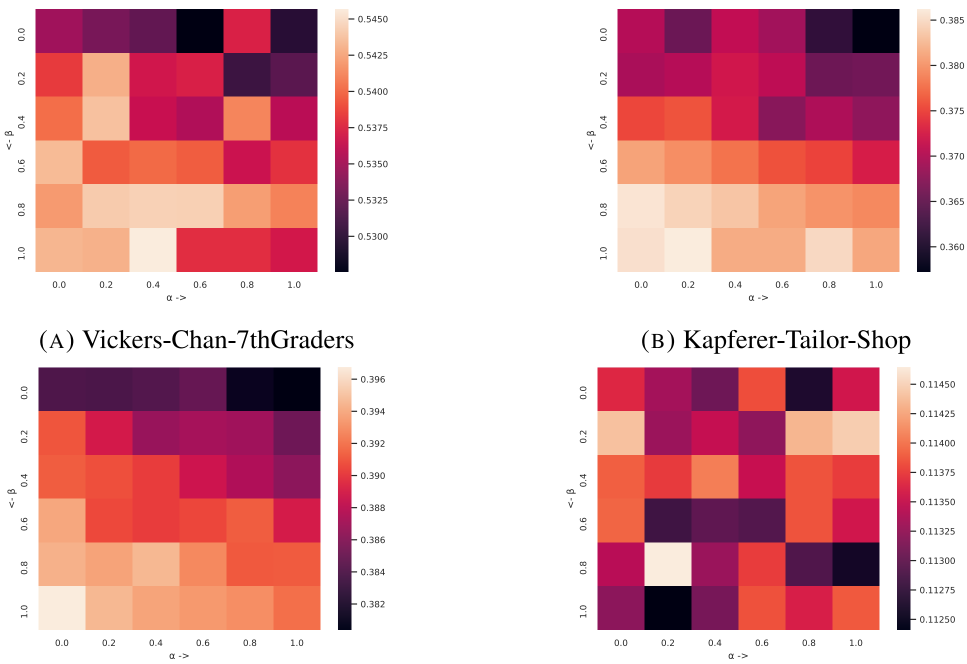


(C) Lazega-Law-Firm



(D) CKM-Physicians-Innovation

FIGURE 5.5: Heatmap of Balanced Accuracy score variation of $MNERLP-MUL$ algorithm's performance with respect to α & β with *Ratio* of testing to total edges averaged over range $0.1-0.5$



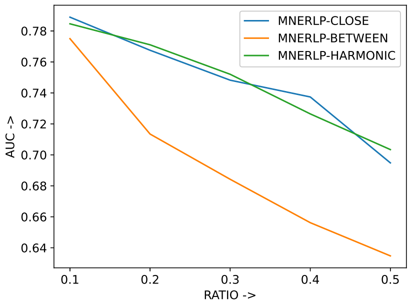
(A) Vickers-Chan-7thGraders (B) Kapferer-Tailor-Shop

(C) Lazega-Law-Firm (D) CKM-Physicians-Innovation

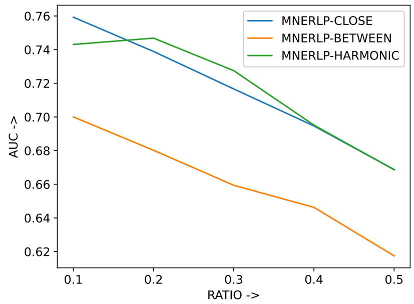
FIGURE 5.6: Heatmap of F1 score variation of $MNERLP-MUL$ algorithm's performance with respect to α & β with *Ratio* of testing to total edges averaged over range $0.1-0.5$

5.3.2 Algorithm Variation Comparison

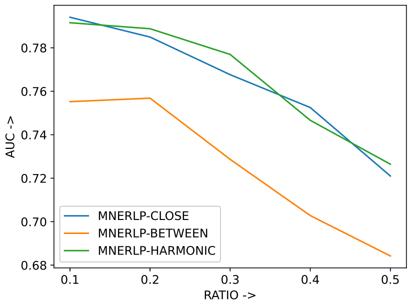
Three different centrality measures have been used (as described in Section 5.1.2) as a benchmark for measuring node relevance. These centrality measures take the entire structure of the graph into account but are not very complex because they use the shortest paths between nodes for calculation purposes. Based on these measures, three possible algorithm variations $MNERLP-MUL$ have been created, which are $MNERLP-CLOSE$, $MNERLP-BETWEEN$, and $MNERLP-HARMONIC$. The relative performance of these variations is compared with each other. The parameters are fixed as $\alpha = 0.2$ and $\beta = 1.0$ based on results of Section 5.3.1.



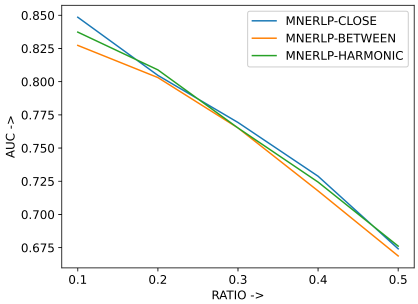
(A) Vickers-Chan-7thGraders



(B) Kapferer-Tailor-Shop



(C) Lazega-Law-Firm

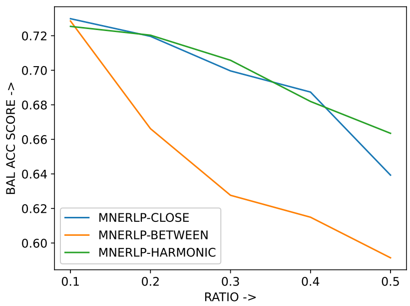


(D) CKM-Physicians-Innovation

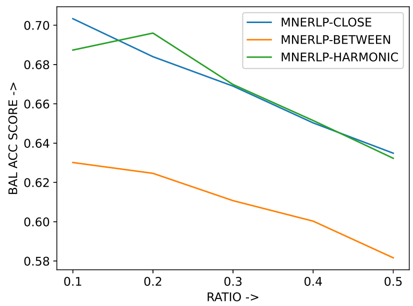
FIGURE 5.7: Graphs of AUC variation of $MNERLP-MUL$ algorithm's performance with respect to different Node Relevance (Centrality) variations

5.3.2.1 Analysis of AUC pattern on different algorithm variations

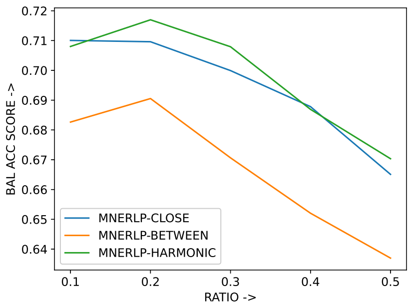
Fig.5.7 compares different algorithm variations on four datasets based on AUC. In all four datasets, it is observed that $MNERLP-CLOSE$ is slightly better than $MNERLP-BETWEEN$ in most cases. However, with the increase in the $Ratio$ variable, the decrease in performance of $MNERLP-CLOSE$ is linear, while in $MNERLP-BETWEEN$, it is more gradual. $MNERLP-HARMONIC$ algorithm has the worst performance in all datasets except in CKM-Physicians-Innovation, where all three variations have comparable performance. The overall pattern is a decrease in AUC values with an increase in the $Ratio$ variable.



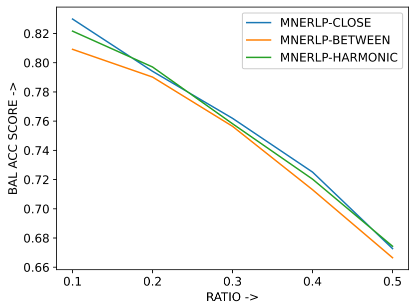
(A) Vickers-Chan-7thGraders



(B) Kapferer-Tailor-Shop



(C) Lazega-Law-Firm

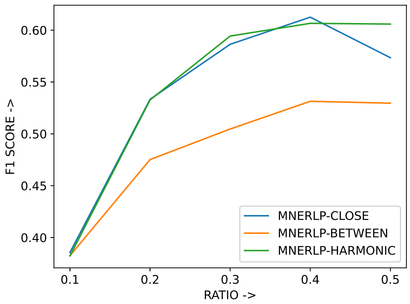


(D) CKM-Physicians-Innovation

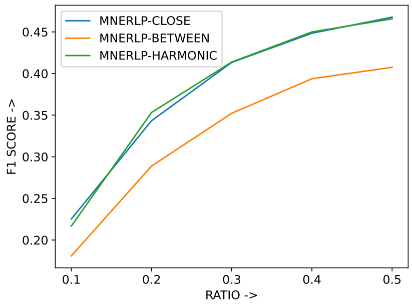
FIGURE 5.8: Graphs of Balanced Accuracy score variation of $MNERLP-MUL$ algorithm's performance with respect to different Node Relevance (Centrality) variations

5.3.2.2 Analysis of Balanced Accuracy score pattern on different algorithm variations

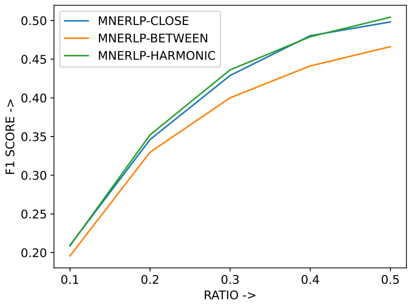
Fig.5.8 compares different algorithm variations on four datasets based on the Balanced Accuracy score. In all four datasets, it is observed that $MNERLP-CLOSE$ is slightly better than $MNERLP-BETWEEN$ in most cases. However, with the increase in the $Ratio$ variable, the decrease in performance of $MNERLP-CLOSE$ is linear, while in $MNERLP-BETWEEN$, it is more gradual. $MNERLP-HARMONIC$ algorithm has the worst performance in all datasets except in CKM-Physicians-Innovation, where all three variations have comparable performance. The overall pattern is a decrease in Balanced Accuracy score values with the increase in the $Ratio$ variable.



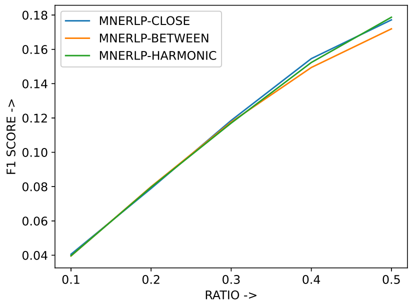
(A) Vickers-Chan-7thGraders



(B) Kapferer-Tailor-Shop



(C) Lazega-Law-Firm



(D) CKM-Physicians-Innovation

FIGURE 5.9: Graphs of F1 score variation of $MNERLP-MUL$ algorithm's performance with respect to different Node Relevance (Centrality) variations

5.3.2.3 Analysis of F1 score pattern on different different variations

Fig.5.9 compares different algorithm variations on four datasets in terms of F1 score. The overall pattern is an increase in F1 score values with the increase in the *Ratio* variable, but the rate of increase decreases gradually. In all four datasets, it is observed that $MNERLP-CLOSE$ and $MNERLP-BETWEEN$ have comparable performance in most cases. $MNERLP-HARMONIC$ algorithm has the worst performance in all datasets except in CKM-Physicians-Innovation, where all three variations have comparable performance. Using the results of this Section 5.3.2, we can use either $MNERLP-CLOSE$ or $MNERLP-BETWEEN$, but $MNERLP-CLOSE$ is chosen because of its relatively better complexity for computation. From henceforth $MNERLP-CLOSE$ with $\alpha = 0.2$ and $\beta = 0.1$ is referred to as $MNERLP-MUL$.

5.3.3 $MNERLP-MUL$ comparison with link prediction methods on summarized weighted graph

Table 5.1 compares the proposed $MNERLP-MUL$ algorithm with baseline methods for the AUC metric. $MNERLP-MUL$ is the best-performing algorithm in all four datasets. The difference is quite drastic for $Ratio$ values in the $0.1-0.3$ range, while for higher $Ratio$ values, $MNERLP-MUL$ is still the best algorithm but with relatively reduced improvement. Overall AUC value decreases with an increase in the $Ratio$ variable. Table 5.2 compares the proposed $MNERLP-MUL$ algorithm with baseline methods for the Balanced Accuracy score. $MNERLP-MUL$ is the best-performing algorithm in all four datasets. The difference is quite drastic for $Ratio$ values in the $0.1-0.3$ range, while for higher $Ratio$ values, $MNERLP-MUL$ is still the best algorithm but with relatively reduced improvement. Overall Balanced Accuracy score value decreases with an increase in the $Ratio$ variable. Table 5.3 compares the proposed $MNERLP-MUL$ algorithm with baseline methods for the F1 score. $MNERLP-MUL$ is the best-performing algorithm in all four datasets. The difference is quite drastic for $Ratio$ values in the $0.3-0.5$ range, while for lower $Ratio$ values, $MNERLP-MUL$ is still the best algorithm but with relatively reduced improvement. Overall F1 score value increases with an increase in the $Ratio$ variable.

5.3.4 $MNERLP-MUL$ comparison with multiplex link prediction methods on individual layers

This section presents the results of the $MNERLP-MUL$ algorithm on application to specific multiplex network layers. Tables 5.4 and 5.5 compare the proposed $MNERLP-MUL$ algorithm with baseline methods for the AUC metric. $MNERLP-MUL$ is the best-performing algorithm in all four datasets across all layers when compared with other weighted link prediction algorithms. The difference in the AUC value between the $MNERLP-MUL$ algorithm and the other baseline weighted

TABLE 5.1: Comparison of the proposed algorithm *MNERLP − MUL* with baseline algorithms in terms of AUC on four datasets and five *Ratio* values for testing to total edges percentage

DATASET	Ratio	CN-WT	JC-WT	PA-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	MNERLP-MUL
Vickers-Chan-7thGraders	0.1	0.60004	0.59855	0.5882	0.6042	0.60601	0.61182	0.59973	0.78879
	0.2	0.59655	0.59499	0.58393	0.60335	0.60082	0.62067	0.59657	0.77604
	0.3	0.5934	0.58616	0.58152	0.59996	0.59932	0.59138	0.58837	0.75642
	0.4	0.59101	0.58617	0.5792	0.59835	0.59769	0.5651	0.58289	0.72095
	0.5	0.59345	0.57801	0.57719	0.59467	0.591	0.54715	0.57768	0.69706
Kapferer-Tailor-Shop	0.1	0.57837	0.57561	0.56307	0.57918	0.58087	0.52853	0.58095	0.75599
	0.2	0.57939	0.56826	0.56007	0.57907	0.58016	0.53631	0.57299	0.73751
	0.3	0.57816	0.56667	0.56138	0.57852	0.58157	0.54292	0.5703	0.71491
	0.4	0.57999	0.56725	0.56024	0.57777	0.58211	0.53724	0.56612	0.69164
	0.5	0.57144	0.56681	0.55963	0.57288	0.57755	0.52632	0.56122	0.65686
Lazega-Law-Firm	0.1	0.58327	0.58782	0.55056	0.58688	0.58607	0.57482	0.58132	0.79472
	0.2	0.58353	0.58566	0.55052	0.58489	0.58376	0.55568	0.57896	0.78692
	0.3	0.58237	0.58298	0.54895	0.58533	0.58243	0.56236	0.57789	0.76985
	0.4	0.582	0.58082	0.5468	0.58114	0.58263	0.55356	0.57243	0.75158
	0.5	0.57966	0.57794	0.54573	0.58069	0.58276	0.53781	0.56917	0.7245
CKM-Physicians-Innovation	0.1	0.66847	0.66582	0.52354	0.67257	0.67278	0.45184	0.68244	0.83587
	0.2	0.65265	0.65394	0.52239	0.65454	0.65404	0.46218	0.67573	0.80663
	0.3	0.6342	0.63574	0.52118	0.63361	0.63369	0.47388	0.66009	0.76035
	0.4	0.61288	0.60942	0.52056	0.61248	0.61329	0.48131	0.64074	0.72673
	0.5	0.59027	0.58902	0.51964	0.58904	0.58998	0.49674	0.61602	0.68041

TABLE 5.2: Comparison of the proposed algorithm $MNERLP - MUL$ with baseline algorithms in terms of Balanced Accuracy score on four datasets and five *Ratio* values

DATASET	Ratio	CN-WT	JC-WT	PA-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	MNERLP-MUL
Vickers-Chan-7thGraders	0.1	0.56597	0.54169	0.55349	0.58436	0.58252	0.5875	0.56001	0.73039
	0.2	0.57452	0.55056	0.55832	0.59176	0.59523	0.59205	0.57174	0.71982
	0.3	0.58677	0.55881	0.57015	0.59978	0.59855	0.56761	0.57577	0.7069
	0.4	0.58636	0.56695	0.56554	0.59867	0.60116	0.54428	0.58033	0.66887
	0.5	0.58643	0.56182	0.57647	0.59716	0.58694	0.53879	0.57449	0.64999
Kapferer-Tailor-Shop	0.1	0.54741	0.53338	0.54904	0.56161	0.56035	0.51314	0.55065	0.69763
	0.2	0.55856	0.54184	0.54904	0.56681	0.56808	0.52591	0.55289	0.68316
	0.3	0.56295	0.54612	0.55579	0.57155	0.57774	0.52316	0.5599	0.67222
	0.4	0.56918	0.55538	0.56164	0.57229	0.58016	0.52915	0.56191	0.64923
	0.5	0.56383	0.55408	0.56106	0.5692	0.57189	0.51929	0.5608	0.61952
Lazega-Law-Firm	0.1	0.56365	0.546	0.53974	0.56838	0.5665	0.55188	0.56111	0.70591
	0.2	0.57049	0.55125	0.54066	0.57218	0.56981	0.54067	0.56485	0.71178
	0.3	0.57152	0.55562	0.54287	0.57509	0.57346	0.5432	0.56894	0.70571
	0.4	0.57516	0.56052	0.54342	0.57435	0.57669	0.54143	0.57057	0.69184
	0.5	0.57158	0.55956	0.54385	0.57647	0.57469	0.52869	0.57314	0.6689
CKM-Physicians-Innovation	0.1	0.66381	0.66194	0.52853	0.66789	0.66808	0.46238	0.6699	0.82019
	0.2	0.64949	0.65105	0.52749	0.65141	0.65071	0.46458	0.66184	0.7954
	0.3	0.63223	0.63411	0.52665	0.63151	0.6316	0.47836	0.65228	0.75365
	0.4	0.6117	0.60865	0.52222	0.61121	0.61214	0.48316	0.63682	0.72282
	0.5	0.58967	0.58875	0.52276	0.58842	0.58938	0.49026	0.61426	0.67879

TABLE 5.3: Comparison of the proposed algorithm *MNERLP − MUL* with baseline algorithms in terms of F1 score on four datasets and five *Ratio* values for testing to total edges percentage

DATASET	Ratio	CN-WT	JC-WT	PA-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	MNERLP-MUL
Vickers-Chan-7thGraders	0.1	0.24939	0.23046	0.23856	0.26631	0.26443	0.26763	0.24456	0.38672
	0.2	0.36801	0.34887	0.35126	0.38635	0.38963	0.3964	0.36634	0.53331
	0.3	0.43959	0.4192	0.42053	0.4525	0.44969	0.44224	0.42935	0.59518
	0.4	0.47248	0.45981	0.45382	0.4817	0.48313	0.45652	0.46641	0.58789
	0.5	0.48319	0.4722	0.48057	0.4904	0.47738	0.48124	0.47389	0.58662
Kapferer-Tailor-Shop	0.1	0.13998	0.13203	0.14099	0.14865	0.14799	0.12286	0.1416	0.22318
	0.2	0.23578	0.22297	0.22674	0.24311	0.24436	0.21383	0.23101	0.34446
	0.3	0.30037	0.28621	0.29205	0.30797	0.31514	0.27168	0.29762	0.41618
	0.4	0.34681	0.33627	0.33903	0.34762	0.35813	0.32355	0.33991	0.44729
	0.5	0.36525	0.36176	0.36781	0.3685	0.37253	0.34197	0.36399	0.44852
Lazega-Law-Firm	0.1	0.1418	0.13149	0.12852	0.14448	0.14351	0.13363	0.14007	0.20773
	0.2	0.23886	0.22318	0.21403	0.2404	0.23829	0.21643	0.23399	0.34766
	0.3	0.3031	0.28959	0.27539	0.3069	0.30494	0.28335	0.30104	0.43474
	0.4	0.35138	0.33946	0.31899	0.34988	0.35287	0.32916	0.3474	0.48389
	0.5	0.37652	0.36942	0.35327	0.38214	0.38008	0.35028	0.38051	0.50254
CKM-Physicians-Innovation	0.1	0.04034	0.04008	0.00759	0.04115	0.04119	0.0053	0.03539	0.03954
	0.2	0.07818	0.07873	0.01487	0.07865	0.07841	0.0105	0.06691	0.07963
	0.3	0.1122	0.11308	0.02206	0.11184	0.11269	0.01678	0.0941	0.11601
	0.4	0.13835	0.13548	0.02836	0.1381	0.13892	0.02258	0.11737	0.15365
	0.5	0.15519	0.15398	0.03522	0.15187	0.15373	0.02885	0.13946	0.18181

algorithms decreases with an increase in the *Ratio* value. In all datasets except CKM-Physicians Innovation, the improvement in performance over other baseline methods is quite drastic. In the CKM-Physicians-Innovation dataset, $LOCALP-WT$ is marginally worse than the $MNERLP-MUL$ method, especially for higher *Ratio* values. In this section, the $MNERLP-MUL$ algorithm is also compared with methods explicitly designed for link prediction in multiplex networks, i.e., $NSILR-MUL$ and $MADM-MUL$. $MNERLP-MUL$ is the best performing algorithm in all three datasets across all layers for *Ratio* values $0.1-0.3$ and primarily for $Ratio = 0.4$. The only notable exception can be seen in layers-1,3 of the Kapferer-Tailor-Shop dataset and layer-1 of the CKM-Physicians-Innovation dataset, where $NSILR-MUL$ and $MADMLP-MUL$ perform better than $MNERLP-MUL$ algorithm. The difference in the AUC value between the $MNERLP-MUL$ algorithm and the other baseline algorithms decreases with an increase in the *Ratio* value for classical weighted link prediction methods. For $NSILR-MUL$ and $MADMLP-MUL$, the pattern of change for increasing *Ratio* value is opposite to that of other algorithms, which gives them a higher performance on fewer edges for any given layer.

Tables 5.6 and 5.7 compare the proposed $MNERLP-MUL$ algorithm with baseline methods for the Balanced Accuracy score. $MNERLP-MUL$ is the best-performing algorithm in all four datasets across all layers. The most minor improvement is observed in the CKM-Physicians-Innovation dataset, similar to the AUC values pattern because of the dataset's structure, which has lower densities and clustering coefficients than other datasets). This also leads to relatively higher average shortest path distances. The overall pattern in both AUC and Balanced Accuracy score metric is a decrease in the performance of algorithms with increasing *Ratio* values. For comparison with $NSILR-MUL$ and $MADMLP-MUL$, the $MNERLP-MUL$ algorithm is better than these baselines for all *Ratio* values $0.1-0.3$ and primarily for $Ratio = 0.4$ in all cases except layer-1 of CKM-Physicians-Innovation and Lazega-Law-Firm and layers-1,3 of Kapferer-Tailor-Shop. The $MNERLP-MUL$ algorithm gives more false negatives than

TABLE 5.4: Comparison of the proposed algorithm *MNERLP – MUL* with baseline algorithms in terms of AUC layer-wise on four datasets and five *Ratio* values for testing to total edges percentage

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Vickers-Chan-7thGraders	1	0.1	0.59264	0.58011	0.59237	0.59553	0.59814	0.57576	0.5941	0.67383	0.64744	0.77473
		0.2	0.59097	0.57682	0.58659	0.59292	0.59469	0.57026	0.58964	0.68865	0.67858	0.75777
		0.3	0.58722	0.57641	0.58255	0.58999	0.59182	0.54356	0.58549	0.67236	0.66006	0.7357
		0.4	0.58726	0.57378	0.57803	0.59082	0.59233	0.53047	0.58239	0.69457	0.67443	0.71536
		0.5	0.58526	0.57038	0.57285	0.58728	0.58691	0.53176	0.57822	0.85148	0.86447	0.67758
	2	0.1	0.61902	0.56741	0.61558	0.6267	0.62889	0.48223	0.59643	0.67613	0.76145	0.84817
		0.2	0.62247	0.56606	0.61424	0.62223	0.63047	0.49103	0.59398	0.70182	0.77866	0.80745
		0.3	0.61905	0.56638	0.60894	0.62182	0.626	0.48118	0.59322	0.7427	0.77261	0.77936
		0.4	0.61329	0.56402	0.61106	0.61451	0.62156	0.50157	0.59178	0.75734	0.79776	0.74621
		0.5	0.60488	0.56501	0.59501	0.60629	0.60729	0.50581	0.58739	0.82937	0.87328	0.71276
	3	0.1	0.62164	0.58049	0.6175	0.62069	0.62662	0.57283	0.60573	0.5735	0.5131	0.83295
		0.2	0.62393	0.57656	0.61901	0.62257	0.62611	0.5692	0.60084	0.61915	0.54899	0.80676
		0.3	0.62152	0.57468	0.61524	0.62353	0.62386	0.5556	0.60198	0.65218	0.63327	0.7902
		0.4	0.62037	0.57429	0.61428	0.62196	0.62074	0.54368	0.6	0.6705	0.66265	0.75691
		0.5	0.61308	0.57382	0.6057	0.6086	0.61482	0.53727	0.60014	0.76442	0.75108	0.72052
Kapferer-Tailor-Shop	1	0.1	0.60486	0.56184	0.59865	0.61264	0.60498	0.52083	0.58524	0.76062	0.80789	0.75131
		0.2	0.60657	0.56166	0.59469	0.60498	0.60136	0.5045	0.58684	0.66902	0.75498	0.73395
		0.3	0.59718	0.5625	0.59257	0.59885	0.59925	0.5177	0.58597	0.76188	0.7278	0.71226
		0.4	0.59247	0.56256	0.58216	0.59263	0.58955	0.52191	0.58082	0.73548	0.74697	0.67976
		0.5	0.58225	0.56422	0.57265	0.57901	0.57752	0.53199	0.57756	0.77726	0.81609	0.65159
	2	0.1	0.5933	0.55929	0.5917	0.59403	0.59351	0.58928	0.57759	0.64462	0.583	0.76397
		0.2	0.59517	0.55953	0.59036	0.5934	0.59279	0.58149	0.5755	0.63217	0.60123	0.74718
		0.3	0.59318	0.55929	0.58669	0.59445	0.59601	0.5807	0.57592	0.62654	0.66526	0.72146
		0.4	0.58828	0.55764	0.58892	0.59235	0.59297	0.5676	0.57596	0.64257	0.68733	0.69211
		0.5	0.5825	0.55684	0.57512	0.58282	0.5831	0.5494	0.57362	0.74455	0.74659	0.66597
	3	0.1	0.59465	0.58391	0.59048	0.59799	0.60194	0.56674	0.58746	0.60643	0.78729	0.70756
		0.2	0.59042	0.57579	0.58732	0.59641	0.58661	0.57093	0.58073	0.71011	0.69995	0.686
		0.3	0.58192	0.5776	0.5823	0.58149	0.58361	0.56323	0.58179	0.65467	0.66214	0.65586
		0.4	0.56753	0.56995	0.57142	0.56861	0.56588	0.56265	0.57163	0.65029	0.6414	0.62762
		0.5	0.55526	0.57049	0.54907	0.55341	0.55506	0.55853	0.55777	0.71653	0.69904	0.59655
	4	0.1	0.606	0.57718	0.59673	0.60136	0.60341	0.5722	0.59919	0.58657	0.51445	0.72213
		0.2	0.60078	0.57634	0.59186	0.59576	0.59168	0.55838	0.60304	0.54893	0.59805	0.6903
		0.3	0.58098	0.57549	0.57476	0.58529	0.58001	0.57257	0.59071	0.57645	0.62804	0.6575
		0.4	0.56782	0.57322	0.56679	0.5688	0.56651	0.56526	0.58244	0.52468	0.61332	0.62506
		0.5	0.55495	0.57306	0.55111	0.558	0.55834	0.54834	0.57537	0.62987	0.65278	0.60746

TABLE 5.5: Comparison of the proposed algorithm $MNERLP-MUL$ with baseline algorithms in terms of AUC layer-wise on four datasets and five *Ratio* values for testing to total edges percentage (contd..)

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Lazega-Law-Firm	1	0.1	0.60144	0.55022	0.60485	0.60372	0.60406	0.53959	0.58762	0.71937	0.77388	0.82854
		0.2	0.60174	0.54922	0.60499	0.60386	0.60404	0.54927	0.58642	0.70427	0.74791	0.80992
		0.3	0.60264	0.54733	0.60502	0.60476	0.60472	0.53716	0.58348	0.72069	0.78127	0.79805
		0.4	0.60464	0.54684	0.60473	0.60476	0.60594	0.53965	0.58262	0.76291	0.82192	0.76883
		0.5	0.60359	0.54577	0.60055	0.60365	0.60216	0.53133	0.58052	0.83277	0.89068	0.74057
	2	0.1	0.64956	0.56805	0.65198	0.65205	0.65521	0.59864	0.61776	0.73276	0.75947	0.8758
		0.2	0.64929	0.56565	0.6512	0.65061	0.65289	0.57781	0.62117	0.72885	0.72875	0.86363
		0.3	0.6468	0.56781	0.64916	0.64993	0.64918	0.56951	0.62255	0.73174	0.74493	0.83866
		0.4	0.64432	0.56695	0.64241	0.64384	0.64376	0.5666	0.62141	0.78454	0.74117	0.81236
		0.5	0.62924	0.56667	0.62691	0.63295	0.63144	0.55269	0.62109	0.83769	0.81413	0.76651
	3	0.1	0.58928	0.54767	0.58937	0.59099	0.5918	0.48343	0.58231	0.58807	0.61789	0.79897
		0.2	0.58961	0.54729	0.58837	0.59161	0.59159	0.50236	0.58153	0.58978	0.59965	0.79021
		0.3	0.59118	0.54629	0.58863	0.59194	0.59218	0.49608	0.57933	0.60929	0.62348	0.76809
		0.4	0.59075	0.54476	0.58684	0.5926	0.59416	0.50003	0.57661	0.61249	0.6409	0.74885
		0.5	0.59025	0.54431	0.58499	0.592	0.59119	0.50672	0.57469	0.74226	0.73718	0.71663
CKM-Physicians-Innovation	1	0.1	0.61102	0.55727	0.60968	0.608	0.61276	0.48	0.64181	0.83846	0.8049	0.72068
		0.2	0.59254	0.55546	0.58915	0.59428	0.59421	0.48811	0.62607	0.82258	0.81245	0.68205
		0.3	0.57691	0.55152	0.57343	0.5777	0.57623	0.5015	0.60527	0.83026	0.8066	0.6448
		0.4	0.55793	0.54855	0.55706	0.55667	0.55735	0.50274	0.58433	0.8176	0.80742	0.6136
		0.5	0.54326	0.54238	0.54133	0.54224	0.54195	0.50455	0.56289	0.81672	0.81927	0.58581
	2	0.1	0.62515	0.52389	0.62195	0.62257	0.62672	0.43557	0.6508	0.69274	0.67307	0.74902
		0.2	0.60593	0.52375	0.60393	0.60433	0.60364	0.45342	0.62561	0.67828	0.70683	0.70204
		0.3	0.58412	0.52232	0.5841	0.58314	0.5837	0.46998	0.60867	0.6835	0.70655	0.66478
		0.4	0.56375	0.52066	0.56396	0.56388	0.56441	0.48797	0.58479	0.68295	0.71381	0.62922
		0.5	0.54599	0.51634	0.54695	0.54776	0.5481	0.48942	0.56361	0.72359	0.74841	0.59098
	3	0.1	0.60595	0.51608	0.60165	0.59935	0.59927	0.47439	0.6197	0.50693	0.58319	0.70811
		0.2	0.58671	0.5171	0.58584	0.58576	0.5842	0.48683	0.60487	0.50875	0.58819	0.66865
		0.3	0.56984	0.51768	0.5694	0.57134	0.57012	0.4896	0.58872	0.5308	0.56816	0.6423
		0.4	0.55439	0.51475	0.55331	0.55466	0.55339	0.49276	0.56967	0.54651	0.58292	0.60516
		0.5	0.53836	0.51305	0.53938	0.53894	0.53906	0.49403	0.55117	0.5848	0.62415	0.57693

TABLE 5.6: Comparison of the proposed algorithm $MNERLP-MUL$ with baseline algorithms in terms of Balanced Accuracy Score layer-wise on four datasets and five *Ratio* values for testing to total edges percentage

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Vickers-Chan-7thGraders	1	0.1	0.56342	0.54828	0.53813	0.56907	0.58082	0.56175	0.55698	0.64623	0.63125	0.7223
		0.2	0.57266	0.55531	0.55013	0.57866	0.59062	0.55696	0.56949	0.63429	0.65912	0.70157
		0.3	0.57415	0.56055	0.55433	0.58016	0.59138	0.53253	0.57322	0.62218	0.63996	0.68246
		0.4	0.57837	0.56688	0.55992	0.58257	0.59203	0.52167	0.5786	0.64011	0.63803	0.66475
		0.5	0.57638	0.56346	0.55444	0.58092	0.58236	0.52659	0.57833	0.77505	0.78723	0.63626
	2	0.1	0.57881	0.54886	0.57224	0.59482	0.61177	0.48568	0.56612	0.55599	0.73898	0.7907
		0.2	0.57762	0.55054	0.57643	0.58365	0.61821	0.48775	0.57738	0.58979	0.75469	0.75206
		0.3	0.58547	0.55241	0.58259	0.58572	0.61037	0.48891	0.58883	0.65324	0.77113	0.73449
		0.4	0.59324	0.5569	0.59702	0.58937	0.59743	0.49477	0.59452	0.69162	0.79962	0.70259
		0.5	0.59323	0.56227	0.59146	0.5928	0.59287	0.49649	0.59123	0.74853	0.85768	0.68686
	3	0.1	0.60595	0.56948	0.58134	0.59791	0.59845	0.56188	0.58746	0.58113	0.51705	0.75878
		0.2	0.5984	0.57069	0.58359	0.60389	0.60532	0.55592	0.58666	0.59704	0.54868	0.74206
		0.3	0.58947	0.57144	0.58588	0.5902	0.60796	0.54277	0.58813	0.64176	0.56173	0.73421
		0.4	0.59494	0.57312	0.59474	0.5935	0.5971	0.53255	0.59372	0.64304	0.60044	0.708
		0.5	0.59621	0.57344	0.59702	0.59039	0.59629	0.51901	0.59698	0.71139	0.64658	0.67095
Kapferer-Tailor-Shop	1	0.1	0.58236	0.57553	0.57355	0.58575	0.59023	0.50667	0.57766	0.73472	0.74219	0.69589
		0.2	0.58235	0.57619	0.57918	0.58307	0.585	0.5036	0.57887	0.63586	0.71668	0.67858
		0.3	0.58075	0.57551	0.57974	0.57988	0.58327	0.50212	0.58127	0.71306	0.69299	0.66833
		0.4	0.58069	0.57333	0.57676	0.57954	0.57812	0.50073	0.57857	0.68368	0.70034	0.65103
		0.5	0.57514	0.5699	0.57148	0.57077	0.57059	0.5142	0.57017	0.70742	0.74115	0.63384
	2	0.1	0.57183	0.55661	0.55885	0.58069	0.58219	0.55825	0.56904	0.60582	0.572	0.70847
		0.2	0.57814	0.56151	0.56307	0.57758	0.58472	0.55377	0.56828	0.55297	0.57243	0.69814
		0.3	0.58139	0.56045	0.56453	0.57695	0.58613	0.55621	0.57678	0.57583	0.63517	0.67729
		0.4	0.5733	0.56157	0.57201	0.57684	0.58162	0.54171	0.5714	0.5935	0.63969	0.65193
		0.5	0.56818	0.56062	0.56647	0.569	0.5724	0.52849	0.56789	0.68637	0.67203	0.63466
	3	0.1	0.58472	0.57213	0.58592	0.58854	0.59169	0.55019	0.5883	0.55734	0.77058	0.67553
		0.2	0.58285	0.56722	0.58491	0.58881	0.57978	0.54079	0.58105	0.66678	0.68634	0.6674
		0.3	0.57704	0.56792	0.58188	0.57669	0.57912	0.53942	0.58105	0.59198	0.63398	0.6438
		0.4	0.56472	0.56328	0.57177	0.56591	0.56394	0.54365	0.56469	0.5756	0.60987	0.62247
		0.5	0.55364	0.56389	0.54944	0.55198	0.55389	0.54634	0.55535	0.67141	0.65416	0.59496
	4	0.1	0.59354	0.57091	0.59038	0.58852	0.59606	0.53106	0.59164	0.58152	0.50236	0.67974
		0.2	0.59011	0.57465	0.59022	0.58578	0.58476	0.52807	0.59539	0.54792	0.58674	0.66767
		0.3	0.57475	0.57473	0.57444	0.57855	0.57422	0.53687	0.57765	0.57568	0.61877	0.65004
		0.4	0.56436	0.57345	0.56694	0.56437	0.56271	0.53096	0.57021	0.52745	0.59535	0.61379
		0.5	0.55324	0.57142	0.55173	0.55582	0.55642	0.53195	0.56193	0.59971	0.59719	0.60131

TABLE 5.7: Comparison of the proposed algorithm $MNERLP-MUL$ with baseline algorithms in terms of Balanced Accuracy Score layer-wise on four datasets and five *Ratio* values for testing to total edges percentage (contd..)

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Lazega-Law-Firm	1	0.1	0.58227	0.54049	0.56931	0.58039	0.58171	0.53054	0.57531	0.64736	0.7537	0.73947
		0.2	0.5819	0.54176	0.57413	0.58313	0.58608	0.53284	0.58075	0.63446	0.73227	0.73132
		0.3	0.58863	0.54207	0.57712	0.5888	0.5879	0.52529	0.58405	0.65072	0.75427	0.72581
		0.4	0.59154	0.54352	0.57722	0.58757	0.58915	0.52991	0.58316	0.6877	0.77753	0.70379
		0.5	0.58151	0.54457	0.57669	0.57972	0.58423	0.52099	0.57887	0.74747	0.81897	0.63831
	2	0.1	0.61336	0.56227	0.61282	0.61364	0.62235	0.56734	0.6119	0.69177	0.75394	0.78236
		0.2	0.61894	0.56044	0.61928	0.61893	0.62182	0.55109	0.61901	0.67491	0.71079	0.78553
		0.3	0.62358	0.56159	0.6254	0.62506	0.62481	0.54644	0.62432	0.67608	0.71223	0.77402
		0.4	0.62817	0.56373	0.62706	0.62642	0.62645	0.53992	0.62546	0.71656	0.68633	0.7561
		0.5	0.61986	0.56247	0.61941	0.62308	0.62191	0.52872	0.62235	0.75864	0.71037	0.73018
	3	0.1	0.57253	0.53898	0.55996	0.57524	0.58014	0.48354	0.5698	0.58468	0.57794	0.72073
		0.2	0.57783	0.54302	0.56308	0.5781	0.58257	0.50276	0.57203	0.57059	0.56017	0.71967
		0.3	0.5778	0.5428	0.5658	0.57998	0.58276	0.49611	0.57696	0.57382	0.58422	0.70474
		0.4	0.58076	0.54152	0.56329	0.58029	0.58334	0.4995	0.57725	0.5843	0.59156	0.69157
		0.5	0.57254	0.54353	0.56452	0.57343	0.57846	0.50193	0.5704	0.68154	0.65967	0.66627
CKM-Physicians-Innovation	1	0.1	0.61021	0.54859	0.60956	0.60696	0.6117	0.46534	0.63944	0.82999	0.80232	0.71719
		0.2	0.59206	0.54731	0.58903	0.59358	0.5935	0.46895	0.62449	0.81372	0.81003	0.67988
		0.3	0.57663	0.54294	0.57341	0.57725	0.57581	0.47447	0.60437	0.8205	0.80408	0.64356
		0.4	0.55777	0.5416	0.55709	0.55643	0.55712	0.48374	0.58381	0.8088	0.80479	0.61295
		0.5	0.54318	0.53711	0.54135	0.54212	0.54183	0.49228	0.56265	0.80729	0.81356	0.58553
	2	0.1	0.62437	0.51267	0.62149	0.62151	0.62569	0.45456	0.64853	0.69206	0.6716	0.7461
		0.2	0.60547	0.51931	0.60368	0.60362	0.60306	0.45975	0.62405	0.67723	0.70542	0.7002
		0.3	0.58381	0.51755	0.58398	0.58273	0.5833	0.46137	0.60766	0.68242	0.70507	0.66383
		0.4	0.56361	0.51815	0.56394	0.56366	0.56421	0.4753	0.58425	0.68217	0.7117	0.62883
		0.5	0.54593	0.51591	0.54694	0.54764	0.54799	0.4841	0.56332	0.72117	0.74452	0.59083
	3	0.1	0.60555	0.51577	0.60115	0.59896	0.59892	0.48115	0.61855	0.50695	0.58152	0.7073
		0.2	0.58653	0.5152	0.58556	0.58553	0.58398	0.49088	0.60413	0.50876	0.58657	0.66848
		0.3	0.56973	0.51436	0.56924	0.57122	0.57	0.49277	0.58833	0.53077	0.56639	0.64241
		0.4	0.55435	0.50925	0.55324	0.5546	0.55335	0.49233	0.56946	0.54643	0.58107	0.60527
		0.5	0.53834	0.50933	0.53935	0.53892	0.53903	0.49378	0.55107	0.58463	0.62221	0.57702

TABLE 5.8: Comparison of the proposed algorithm $MNERLP-MUL$ with baseline algorithms in terms of F1 Score layer-wise on four datasets and five *Ratio* values for testing to total edges percentage

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Vickers-Chan-7thGraders	1	0.1	0.23234	0.2198	0.21341	0.23703	0.24751	0.23074	0.22711	0.03538	0.03269	0.35665
		0.2	0.35137	0.33263	0.33206	0.35722	0.36902	0.343	0.34857	0.08204	0.08616	0.49255
		0.3	0.41243	0.39766	0.39899	0.41899	0.42884	0.38755	0.41402	0.16237	0.16881	0.55021
		0.4	0.4486	0.44206	0.44207	0.45384	0.46286	0.41806	0.45155	0.32651	0.32604	0.57228
		0.5	0.46538	0.45956	0.4552	0.46777	0.46607	0.45903	0.47055	0.73093	0.74251	0.56238
	2	0.1	0.10942	0.09536	0.10508	0.11881	0.1306	0.07326	0.10219	0.01307	0.02225	0.19832
		0.2	0.18623	0.1666	0.18549	0.19162	0.22908	0.13247	0.18591	0.03656	0.05672	0.31163
		0.3	0.24901	0.21787	0.24611	0.24902	0.27958	0.17645	0.25216	0.10315	0.1365	0.38509
		0.4	0.30354	0.26385	0.30826	0.29892	0.30889	0.21908	0.30523	0.22518	0.29543	0.41986
		0.5	0.33242	0.30017	0.33042	0.33154	0.33195	0.24781	0.32988	0.50117	0.65207	0.4482
	3	0.1	0.16091	0.13369	0.14028	0.15392	0.15346	0.12536	0.14546	0.02887	0.02294	0.22638
		0.2	0.24395	0.21794	0.22878	0.25072	0.25272	0.20568	0.23238	0.07093	0.0517	0.35641
		0.3	0.29675	0.27762	0.29286	0.29755	0.31927	0.25776	0.29526	0.15756	0.10472	0.4424
		0.4	0.34714	0.32164	0.34699	0.34529	0.34926	0.29383	0.34572	0.24787	0.23321	0.47667
		0.5	0.37711	0.3533	0.37783	0.36839	0.37606	0.31429	0.37774	0.52324	0.44641	0.47788
Kapferer-Tailor-Shop	1	0.1	0.07427	0.07089	0.06992	0.07494	0.07898	0.04998	0.07161	0.02082	0.01848	0.10392
		0.2	0.13276	0.12813	0.13118	0.13418	0.13811	0.09055	0.13085	0.03534	0.04003	0.17658
		0.3	0.1821	0.17427	0.18098	0.18125	0.18565	0.1264	0.18262	0.09088	0.078	0.23799
		0.4	0.22294	0.21052	0.21849	0.22174	0.2205	0.15428	0.22031	0.16412	0.16132	0.27476
		0.5	0.24558	0.23636	0.24025	0.23927	0.23913	0.18857	0.23836	0.35341	0.37883	0.29652
	2	0.1	0.10429	0.09707	0.09689	0.10982	0.11103	0.09363	0.10329	0.02261	0.01838	0.15579
		0.2	0.18211	0.16857	0.16894	0.18215	0.18897	0.16093	0.17428	0.0395	0.04457	0.25638
		0.3	0.24112	0.21989	0.22364	0.23603	0.24636	0.21696	0.23556	0.08119	0.11703	0.31941
		0.4	0.27473	0.26253	0.27378	0.27883	0.28436	0.25129	0.27277	0.16636	0.22874	0.3544
		0.5	0.29691	0.29252	0.29456	0.29761	0.30121	0.27141	0.29675	0.43194	0.45028	0.37969
	3	0.1	0.04704	0.03482	0.04729	0.04797	0.04912	0.02818	0.04673	0.00743	0.02184	0.04769
		0.2	0.09182	0.06439	0.09283	0.09525	0.08897	0.05295	0.08622	0.01833	0.02508	0.09506
		0.3	0.12372	0.08904	0.12875	0.12325	0.12518	0.0725	0.11969	0.03298	0.04533	0.12924
		0.4	0.14541	0.11239	0.15598	0.14706	0.14531	0.09835	0.13714	0.06615	0.08715	0.16499
		0.5	0.15248	0.13211	0.14472	0.14988	0.15243	0.11997	0.14278	0.20833	0.22034	0.17833
	4	0.1	0.0539	0.04248	0.05299	0.05274	0.05737	0.03234	0.05298	0.01487	0.00606	0.06339
		0.2	0.09755	0.07811	0.0977	0.09455	0.09648	0.05925	0.09932	0.02495	0.03786	0.11421
		0.3	0.13081	0.11299	0.13029	0.13382	0.1308	0.08876	0.13106	0.05896	0.0814	0.15361
		0.4	0.15224	0.13955	0.1548	0.15124	0.15027	0.10911	0.15578	0.06382	0.11305	0.16701
		0.5	0.16357	0.16545	0.16018	0.16728	0.16809	0.12896	0.17322	0.20017	0.20173	0.19778

TABLE 5.9: Comparison of the proposed algorithm $MNERLP-MUL$ with baseline algorithms in terms of F1 Score layer-wise on four datasets and five *Ratio* values for testing to total edges percentage (contd..)

DATASET	Layer No.	Ratio	CN-WT	PA-WT	JC-WT	AA-WT	RA-WT	CC-WT	LOCALP-WT	NSILR-MUL	MADMLP-MUL	MNERLP-MUL
Lazega-Law-Firm	1	0.1	0.10214	0.0841	0.09518	0.10104	0.10174	0.07993	0.09834	0.01899	0.02315	0.15181
		0.2	0.17668	0.149	0.17004	0.17762	0.18031	0.14416	0.17554	0.04393	0.05407	0.26168
		0.3	0.24073	0.19951	0.22858	0.24073	0.23982	0.19004	0.23559	0.09293	0.11807	0.34939
		0.4	0.28788	0.24043	0.27224	0.28322	0.28512	0.23379	0.27832	0.2059	0.26308	0.38917
		0.5	0.31087	0.27351	0.30575	0.30886	0.3132	0.26031	0.30794	0.47663	0.57358	0.41974
	2	0.1	0.06509	0.04836	0.06448	0.06511	0.06906	0.04634	0.06457	0.01364	0.01554	0.09425
		0.2	0.1246	0.08941	0.12418	0.12486	0.12707	0.08246	0.12428	0.03097	0.03315	0.17808
		0.3	0.18069	0.12563	0.18083	0.18098	0.18119	0.11466	0.18004	0.0646	0.06911	0.24708
		0.4	0.23264	0.15851	0.23109	0.23088	0.23147	0.14149	0.22904	0.14573	0.1355	0.29577
		0.5	0.26831	0.18504	0.26787	0.27274	0.2718	0.16068	0.27041	0.32698	0.31678	0.32051
	3	0.1	0.09872	0.08486	0.09263	0.10014	0.10288	0.06751	0.09718	0.01853	0.03081	0.15035
		0.2	0.1762	0.15191	0.16469	0.17621	0.18034	0.13095	0.17129	0.04217	0.05624	0.26132
		0.3	0.23195	0.20213	0.22089	0.2341	0.23718	0.17266	0.23096	0.07876	0.10205	0.33291
		0.4	0.27875	0.24122	0.26153	0.27823	0.28158	0.21257	0.27511	0.15454	0.18741	0.38402
		0.5	0.30293	0.27502	0.29498	0.30403	0.30829	0.24412	0.30097	0.41764	0.41913	0.40569
CKM-Physicians-Innovation	1	0.1	0.03032	0.00422	0.03018	0.02945	0.03054	0.00245	0.02253	0.00469	0.01348	0.02887
		0.2	0.05594	0.00836	0.05369	0.05656	0.0566	0.005	0.04459	0.01061	0.03222	0.05448
		0.3	0.07733	0.01228	0.07431	0.0774	0.07641	0.00765	0.06515	0.02313	0.06463	0.07545
		0.4	0.08621	0.01635	0.08542	0.08468	0.08597	0.0105	0.08118	0.04804	0.12515	0.09292
		0.5	0.08888	0.01989	0.08535	0.08716	0.08667	0.0135	0.09094	0.11108	0.23608	0.10391
	2	0.1	0.03712	0.00375	0.03641	0.03645	0.03754	0.00259	0.02748	0.00617	0.01295	0.03544
		0.2	0.06896	0.0078	0.06801	0.06791	0.06783	0.00526	0.051	0.01319	0.03482	0.06585
		0.3	0.09024	0.0115	0.0902	0.08921	0.09026	0.00794	0.07458	0.02801	0.06779	0.09216
		0.4	0.10079	0.01536	0.10104	0.10124	0.10208	0.01108	0.09025	0.05751	0.12701	0.11291
		0.5	0.09887	0.01875	0.10029	0.10193	0.10231	0.01398	0.09755	0.13954	0.23845	0.11668
	3	0.1	0.04092	0.00329	0.03949	0.03878	0.03858	0.00262	0.0297	0.0023	0.01399	0.0379
		0.2	0.0708	0.00647	0.07033	0.07007	0.06925	0.0055	0.0566	0.00602	0.03415	0.06813
		0.3	0.09088	0.0096	0.09095	0.09304	0.09161	0.00822	0.07999	0.03062	0.04758	0.09613
		0.4	0.10077	0.01237	0.09908	0.10111	0.09921	0.01071	0.09435	0.06182	0.08763	0.10841
		0.5	0.09291	0.01538	0.09506	0.09399	0.09442	0.01289	0.0982	0.12961	0.17138	0.11241

other link prediction algorithms designed for link prediction in multiplex networks (for 0.5 probability threshold).

Tables 5.8 and 5.9 compare the proposed $MNERLP - MUL$ algorithm with baseline methods for the F1 score. $MNERLP - MUL$ is the best-performing algorithm in all four datasets across all layers when compared with weighted link prediction methods. The improvement in the three datasets is quite drastic, with CKM-Physicians-Innovation being the exception where for higher $Ratio$ values, the performance of the $MNERLP - MUL$ algorithm is only slightly better than other baseline methods. Compared with $NSILR - MUL$ and $MADMLP - MUL$, the $MNERLP - MUL$ algorithm is better than these baselines for all $Ratio$ values $0.1 - 0.4$. The $Ratio = 0.5$ $MNERLP - MUL$ algorithm performs slightly worse than the algorithm designed specifically for link prediction in muiltiplex networks. Contrary to the pattern in AUC, all three $NSILR - MUL$, $MADMLP - MUL$, and $MNERLP - MUL$ display an improved performance with an increase in the $Ratio$ variable.

5.4 Concluding Remarks

In this chapter, a novel method for link prediction in multiplex networks is presented ($MNERLP - MUL$) based on merging node and edge relevance to take both local and global information into account. The proposal aimed to predict links using more information between nodes (quasi-local approach) and to better predict links in specific layers from a summarized weighted graph. The edge relevance of existing edges is based on local information (3-Degree of Influence), while node relevance is based on global information (centrality measure). The relative contributions of these factors for best link prediction performance are also explored. The results demonstrate that local neighborhood-based algorithms take a very restrained view of comprehensive information to predict edges between nodes, resulting in lower accuracy. We have improved upon this fact. $MNERLP - MUL$ approach is different from standard approaches to link prediction in multiplex networks because an approach with two characteristics has been developed. First, it can be applied to any weighted graphs (summarized from a full multiplex graph), and the results can be directly used for link prediction in all layers of the multiplex network. Another characteristic is that only one round of link prediction should be performed (non-layer specific). Layer-specific link likelihoods can be calculated with just a simple multiplication with an unpacking constant. $MNERLP-MUL$ method outperforms baseline methods used for link prediction in weighted static networks (including $LOCALP-WT$, a quasi-local method) and link prediction methods designed for link prediction in multiplex networks ($NSILR-MUL$ and $MADMLP-MUL$).