Gating-enhanced hierarchical structure learning in hyperbolic space and multi-scale neighbor topology learning in Euclidean space for prediction of microbe-drug associations

Ping Xuan,[†,‡] Chunhong Guan,[†] Sentao Chen,[‡] Jing Gu,[†] Xiuju Wang,[†] Toshiya Nakaguchi,[¶] and Tiangang Zhang[*,†,§]

†*School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China*

‡*Department of Computer Science and Technology, Shantou University, Shantou 515063, China*

¶*Center for Frontier Medical Engineering, Chiba University, Chiba 2638522, Japan*

§*School of Mathematical Science, Heilongjiang University, Harbin 150080, China*

E-mail: zhang@hlju.edu.cn

Abstract

Identifying drug-related microbes may help explore how the microbes affect the functions of drugs by promoting or inhibiting their effects. Most previous methods for prediction of microbe-drug associations focused on integrating the attributes and topologies of microbe and drug nodes in Euclidean space. The heterogeneous network composed of microbes and drugs has hierarchical structure and the hyperbolic space is helpful for reflecting the structure. However, the previous methods did not fully exploit the structure. We propose a prediction method, MFLP, to fuse the hierarchical structure of microbe and drug nodes in hyperbolic space and the multi-scale neighbor topologies in Euclidean space. First, we project the nodes of the

microbe-drug heterogeneous network on the sphere in hyperbolic space, and then construct a topology which implies hierarchical structure and form a hierarchical attribute embedding. The node information from multiple types of neighbor nodes with the new topological structure in the tangent plane space of sphere is aggregated by the designed gating-enhanced hyperbolic graph neural network. Second, the gate at the node feature level is constructed to adaptively fuse the hierarchical features of microbe and drug nodes from two adjacent graph neural encoding layers. Third, multiple neighbor topological embeddings for each microbe and drug node are formed by neighborhood random walks on the microbe-drug heterogeneous network, and they cover neighborhood topologies with multiple scales respectively. Finally, as each scale of topological embedding contains its specific neighborhood topology, we establish an independent graph convolutional neural network for the topology and form the topological representations of microbe and drug nodes in Euclidean space. The comparison experiments based on cross validation showed MFLP outperformed than several advanced prediction methods, and the ablation experiments verified the effectiveness of MFLP's major innovations. The case studies on three drugs further demonstrated MFLP's ability in being applied to discover potential candidate microbes for the given drugs.

Introduction

Microorganisms, including bacteria, fungi, viruses, and protozoa, are organisms that cannot be seen with the naked eye.[1–3] They are present in the oral cavities and intestinal tissues of humans.[4–6] Imbalances and disruptions in these microorganisms can easily lead to various diseases, such as periodontal diseases,[7,8] diabetes,[9,10] inflammatory bowel disease,[11,12] esophageal cancer,[13,14] and anogenital cancers.[15] Antibiotics, like penicillin, are used to treat several diseases caused by microbial infections.[16] However, interactions between microbes and drugs can result in antibiotic resistance in humans.[17–20] Therefore, it is crucial to identify microbe-drug associations in order to explore microbial resistance mechanisms and facilitate research on new drugs.

Computational methods can predict microbial candidates associated with drugs and filter reliable associations between microbes and drugs for subsequent biological experiments. Zhu *et al.*[21]constructed a microbe-drug heterogeneous network based on microbe

similarity, drug similarity, and known microbe-drug associations. They utilized the KATZ method to calculate the correlation of microbe (drug) nodes in the network. HNERMDA[22] incorporated metapath2vec to learn the low-dimensional features of microbe and drug nodes and then infer the tendency of their association. Zhu *et al.*[23] also proposed a computational model based on the Laplacian regularization least-squares method. However, these prediction models are limited and do not excavate deeper connections between microbes and drugs.

Recently, deep learning has demonstrated excellent capability in extracting deep features of microbes and drugs, thereby enhancing prediction performance. For example, a prediction method was proposed that combines conditional random field and graph convolutional network to obtain node representations with improved similarity.[24] Long *et al.*[25] proposed a graph neural network with node-level and graph-level attention to learn embeddings of microbe and drug nodes based on multiple microbe-drug networks. GSAMDA, a prediction model based on a sparse autoencoder and graph attention network autoencoder, learns node topology features and attribute features separately.[26] The Graph2MDA method built a graph variational encoder to learn the feature distribution of microbe-drug node pairs for microbe-drug prediction.[27] Xuan *et al.*[28] designed relational-aware graph reasoning to predict drug-associated microbes. However, these methods were designed for Euclidean spaces. The hyperbolic space contains information about the hierarchical structure of the microbe-drug network and this information cannot be adequately captured by these methods.

We propose a **m**ulti-space **f**eature **l**earning enhanced microbe-drug association **p**rediction method, MFLP, to encode and fuse hierarchical structure of microbe and drug nodes in hyperbolic space and the neighbor topologies of a target node in Euclidean space. The contributions of our method are listed as following.

First, the hyperbolic space may help to reflect and encode the hierarchical structure composed of microbe and drug nodes. Most of the previous prediction models were established which utilized the topology and attribute in Euclidean space, and they ignored the hierarchical structure composed of drugs and microbes. The original features of drugs and microbes were projected to hyperbolic space and the projected features implied more

hierarchical information of two types of nodes.

Second, as multiple drug (microbe) neighbors have different importance for a target drug (microbe) node, we designed a gating-enhanced hyperbolic graph neural network (GHGNN) to integrate the attributes of neighbors on the sphere. A feature-level gate is constructed to adaptively integrate the hierarchical features of each drug (microbe) node of the current GHGNN encoding layer and those of the previous encoding layer. It is helpful for relieving the over-smoothing problem of hyperbolic graph neural network.

Third, the node feature learning in hyperbolic space focused on revealing the hierarchical structures composed of multiple types of nodes, while Euclidean space is beneficial for the neighbor topology learning for each node in the heterogeneous network. We constructed multiple topological embeddings of each drug (microbe) by neighborhood random walks on the microbe-drug heterogeneous network. These embeddings covered the neighbor topologies with multiple scales, and each embedding has its specific topological features.

Finally, for each embedding, an individual graph convolutional neural network was constructed to fuse its specific topology and the attributes of drug and microbe nodes. The contributions of the node features of multiple scales were discriminated and these features were fused adaptively to form the comprehensive neighborhood topological representation. The comparison with six advanced methods and the case studies on three drugs showed the superior prediction accuracy and the ability in discovering the potential candidate drug-microbe associations.

Materials and methods

To predict the microbes associated with a given drug, we propose a microbe-drug association prediction model, MFLP, based on data constructed from drug similarity, microbe similarity, and microbe-drug associations (Fig.1). Initially, we established a heterogeneous network to represent the similarities and associations between drugs and microbes. Microbe nodes and drug nodes hierarchical structure features in hyperbolic space and node multi-scale topological features in Euclidean space are encoded and learned separately. Finally, the hyperbolic hierarchical structure features and multi-scale topological features are integrated by a multi-layer perceptron. The higher the predicted value of the

microbe-drug association, the more likely the association is. The predicted value of the microbe-drug association indicates the level of association, the higher the predicted value, the greater the relevance.

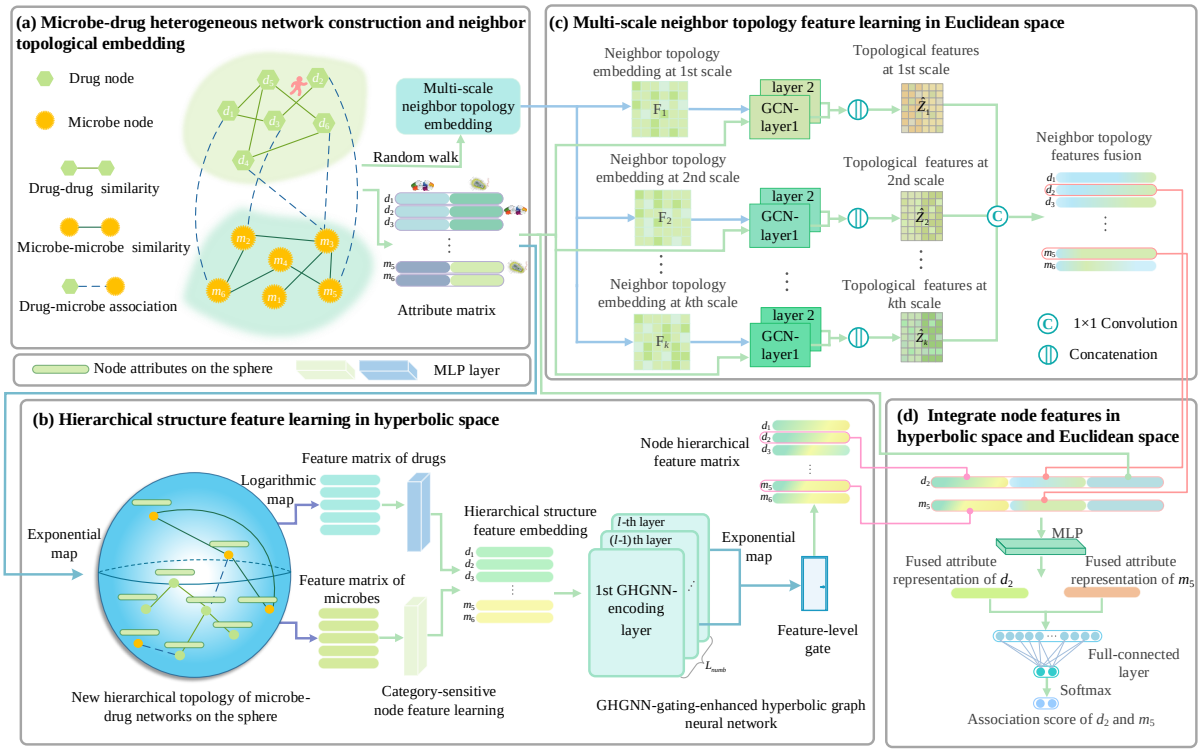


Figure 1: Overview of the proposed MFLP model. **(a)** construct a microbe-drug heterogeneous network and form the neighbor topological embedding by random walks. **(b)**learn hierarchical structure features in hyperbolic space by GHGNN. **(c)** learn neighbor topology features with multiple scales in Euclidean space. **(d)** integrate hierarchical structure features and multi-scale neighbor topology features and estimate the association score of a pair of drug and microbe nodes.

Dataset

In a previous study,[25] we collected data regarding drug similarity, microbe similarity, and microbial-drug associations. We extracted a total of 2,470 known pairs of microbe-drug associations from the Microbe-Drug Association Database (MDAD), which encompassed 1,373 drugs and 173 microbes.[29] To calculate the Gaussian kernel similarity between drugs, known drug interactions were used. Additionally, according to the medicinal chemical substructures, the SIMCOMP2 method was employed to obtain structural similarity scores of drugs.[30] The microbial functional similarity was calculated using the link scores of nodes in the protein-microbe network.

Construction of microbe-drug bi-layer heterogeneous network and feature denotation

We establish a microbe-drug bi-layer heterogeneous network $G = (U, T, H)$ based on drug similarity, microbe similarity, and microbe-drug association. The node set U includes drug nodes U^{drug} and microbe nodes U^{micr}, with an edge $t_{ij} \in T$ connecting a pair of nodes $u_i, u_j \in U$ with a weight of $h_{ij} \in H$. The weight matrix of edges is defined as $H = (M, B)$, which includes the association matrix M containing nodes of different types and the similarity matrix B containing nodes of the same type. The association matrix M with different nodes is defined as,

$$M \in R^{N_d \times N_m},\ if\ u_i \in U^{drug},\ u_j \in U^{micr} \tag{1}$$

where N_d and N_m are the number of drug nodes and microbe nodes, respectively. If there is an association between the drug u_i and the microbe u_j , the M_{ij} value is 1; otherwise, it is 0. The association matrix B for the same nodes is defined as,

$$B = \begin{cases} B^{drug} = (B_{ij}^{drug}) \in R^{N_d \times N_d},\ if\ u_i,\ u_j \in U^{drug} \\ B^{micr} = (B_{ij}^{micr}) \in R^{N_m \times N_m},\ if\ u_i,\ u_j \in U^{micr} \end{cases} \tag{2}$$

where B^{drug} is the drug similarity matrix and B^{micr} is the microbe similarity matrix. The values of $B_{ij}^{drug}(B_{ij}^{micr})$ range from $[0, 1]$, with the higher value indicating higher similarity between the two drugs (microbes). Based on the association matrix of different nodes and the similarity matrix of similar nodes, the drug-microbe adjacency matrix is defined as,

$$S = \begin{bmatrix} B^{drug} & M \\ M^T & B^{micr} \end{bmatrix} \tag{3}$$

where $S \in R^{N_a \times N_a}(N_a = N_d + N_m)$ and M^T is the transpose of the association matrix M. The i-th row of S , denoted as S_i , contains the associations of the drug (microbe) node $u_i \in U$ with other drug (microbe) nodes and the similarities with all microbe (drug) nodes. Therefore, S_i can be used as an attribute vector for node i. Consequently, S_i can serve as a node attribute matrix in the heterogeneous network and is renamed X_{feat}.

Learning hierarchical structure features in hyperbolic space

The microbe-drug heterogeneous network contains similarities and associations among multiple nodes. Upon calculating the δ-hyperbolicity value,[31,32] the microbe-drug network registers a value of 1.5, indicating a pronounced hierarchical structure within it. Prior research[33–35] has highlighted that networks exhibiting hierarchical structures are better suited for modeling in hyperbolic spaces. Hence, we build a gating-enhanced hyperbolic graph neural network module to learn hierarchical structure features that contain information about the inherent hierarchy of the network.

Feature mapping of drug (microbe) nodes in hyperbolic space

Given the feature matrices X_{feat} of drugs and microbes, consider drug u_i as an example. The initial feature vector X^i_{feat} of u_i is assumed to reside in Euclidean space. To render the node features accessible in hyperbolic space, we use $exp^c_x(\cdot)$ exponential mapping to project them into the Poincaré ball of the hyperbolic space.

$$\exp_x^c\left(X^i_{\text{feat}}\right) = x \oplus_c \left(\tanh\left(\sqrt{|c|}\frac{\lambda_x^c \left\|X^i_{\text{feat}}\right\|}{2}\right)\frac{X^i_{\text{feat}}}{\sqrt{|c|}\left\|X^i_{\text{feat}}\right\|}\right) \tag{4}$$

$P^n_c = \{\eta \in \mathbb{R}^n : \|\eta\|^2 < \frac{1}{c}\}$ is an open n-dimensional ball with radius $\frac{1}{\sqrt{c}}$. x is positioned at a point on the hyperbolic space, $\oplus_c$ represents Mobius addition, (further elaborated in Eq. 10). $\lambda^c_x = \frac{2}{1-c\|x\|^2}$ is a conformal factor, and $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ is the Euclidean distance from the point $x \in R^{1\times N_a}$ to the origin. $\partial_x P^d_c$ denotes the tangent space of point x on the Poincaré ball. Assuming X^i_{feat} lies in the tangent space $\partial_0 P^d_c$ of point $x = 0$, we can obtain the feature vector H^i_{feat} of drug u_i positioned on the Poincaré sphere,

$$H^i_{feat} = \exp_0^c\left(X^i_{\text{feat}}\right) \tag{5}$$

All microbe nodes and drug nodes in the heterogeneous network G are projected into the hyperbolic space using exponential mapping, forming a feature matrix $H_{feat} \in R^{N_a \times N_a}$.

Establishing hierarchical structure embedding of nodes

The matrix formed by the rows corresponding to the first N_d nodes of H_{feat} is the hyperbolic feature H^{drug}_{feat} matrix of all drugs, and the matrix formed by the $N_d + 1$ to N_a feature vectors is the hyperbolic feature matrix H^{micr}_{feat} of all microbes. Since addition

and multiplication of drug (microbe) node feature vectors cannot be directly performed on Poincaré spheres, tangent plane feature matrices $\log_x^c(H_{feat}^{drug})$ and $\log_x^c(H_{feat}^{micr})$ can be obtained from the hyperbolic feature matrix H_{feat}^{drug} of all drug nodes and the hyperbolic feature matrix H_{feat}^{micr} of microbe nodes by logarithmic mapping,

$$\log_x^c(H_{feat}^{drug}) = \frac{2}{\sqrt{|c|}\lambda_x^c} \tanh^{-1}(\sqrt{|c|}\left\|-x \oplus_c H_{feat}^{drug}\right\|_2) \frac{\left\|-x \oplus_c H_{feat}^{drug}\right\|}{\left\|-x \oplus_c H_{feat}^{drug}\right\|_2} \tag{6}$$

$$\log_x^c(H_{feat}^{micr}) = \frac{2}{\sqrt{|c|}\lambda_x^c} \tanh^{-1}(\sqrt{|c|}\left\|-x \oplus_c H_{feat}^{micr}\right\|_2) \frac{\left\|-x \oplus_c H_{feat}^{micr}\right\|}{\left\|-x \oplus_c H_{feat}^{micr}\right\|_2} \tag{7}$$

There are two types of nodes in the hyperbolic space, the microbe and drug nodes, so there is node heterogeneity. We designed a category-sensitive node feature learning strategy to obtain the drug (microbe) node hierarchy attributes and form $\tilde{H}_{feat}^{drug}(\tilde{H}_{feat}^{micr})$.

$$\tilde{H}_{feat}^{drug} = (\exp_x^c(W^{H,drug} \otimes_c \log_x^c(H_{feat}^{drug}))) \oplus_c (\exp_x^c(\delta^{drug})) \tag{8}$$

$$\tilde{H}_{feat}^{micr} = (\exp_x^c(W^{H,micr} \otimes_c \log_x^c(H_{feat}^{micr}))) \oplus_c (\exp_x^c(\delta^{micr})) \tag{9}$$

where $W^{H,drug} \in R^{N_d \times d}$ and $W^{H,micr} \in R^{N_m \times d}$ are transformation matrices for the drugs and microbes. δ^{drug} and δ^{micr} are the biases of the tangent space, and $\otimes_c$ represents Mobius matrix-vector multiplication. Taking $\tilde{H}_{feat}^{drug}$ as an example and denoting $\exp_x^c(W^{H,drug} \otimes_c \log_x^c(H_{feat}^{drug}))$ and $\exp_x^c(\delta^{drug})$ as τ_i and τ_j, respectively, the operator $\oplus_c$ is Mobius addition in P_c^d as follows:

$$\tau_i \oplus_c \tau_j := \frac{(1 + 2c\langle\tau_i, \tau_j\rangle + c\|\tau_i\|^2)\tau_i + (1 - c\|\tau_i\|^2)\tau_j}{1 + 2c\langle\tau_i, \tau_j\rangle c^2\|\tau_i\|^2\|\tau_j\|^2} \tag{10}$$

Finally, $\tilde{H}_{feat}^{drug}$ and $\tilde{H}_{feat}^{micr}$ are stacked vertically to form all node hierarchical attribute embeddings $\tilde{H}_{feat}$.

Learning hierarchical structure features by a gating-enhanced hyperbolic graph neural network

Microbes and drugs exhibit connections of varying strengths in the heterogeneous network G, which significantly contributes to prediction accuracy. Thus, we propose a gating-enhanced hyperbolic graph neural network to learn the hierarchical structure features of

microbe (or drug) nodes (Fig.2). The importance of the microbe (or drug) neighbor node u_j to the target node u_i at the l-th level is normalized as Θ_{ij}^{l},

$$\Theta_{\mathrm{ij}}^{l}=\frac{\rho\left(ELU\left(q^{T}\left[\log_{x}^{c}\left(\tilde{H}_{\mathrm{feat}}^{i,l-1}\right)\|\log_{x}^{c}\left(\tilde{H}_{\mathrm{feat}}^{j,l-1}\right)\right]\right)\right)}{\sum_{a\in N_{(i)}}\rho\left(ELU\left(q^{T}\left[\log_{x}^{c}\left(\tilde{H}_{\mathrm{feat}}^{i,l-1}\right)\|\log_{x}^{c}\left(\tilde{H}_{\mathrm{feat}}^{i,l-1}\right)\right]\right)\right)} \tag{11}$$

where q^T is the weight vector, and $\rho(\cdot)$ is the exponential function. $N(i)$ represents all the neighbors of the drug (microbe) node u_i. "||" denotes the concatenation operation, and $ELU(\cdot)$ is a nonlinear activation function ELU.[36] $l \in [1, L_{numb}]$ and L_{numb} is the number of layers of a hyperbolic heterograph neural network, and $\tilde{H}_{feat}^{i,0} = \tilde{H}_{feat}^{i}$ when l = 1. Through neighbor aggregation, the hierarchical structure features of the l-th level of u_i are represented as:

$$Z_{feat}^{i,(l)} = ReLU(\exp_{x}^{c}(\sum_{j\in N(i)} \Theta_{ij}^{l} \cdot \log_{x}^{c}(\tilde{H}_{feat}^{j,(l-1)}))) \tag{12}$$

where $ReLU(\cdot)$ is the nonlinear activation function ReLU. Considering that an increase

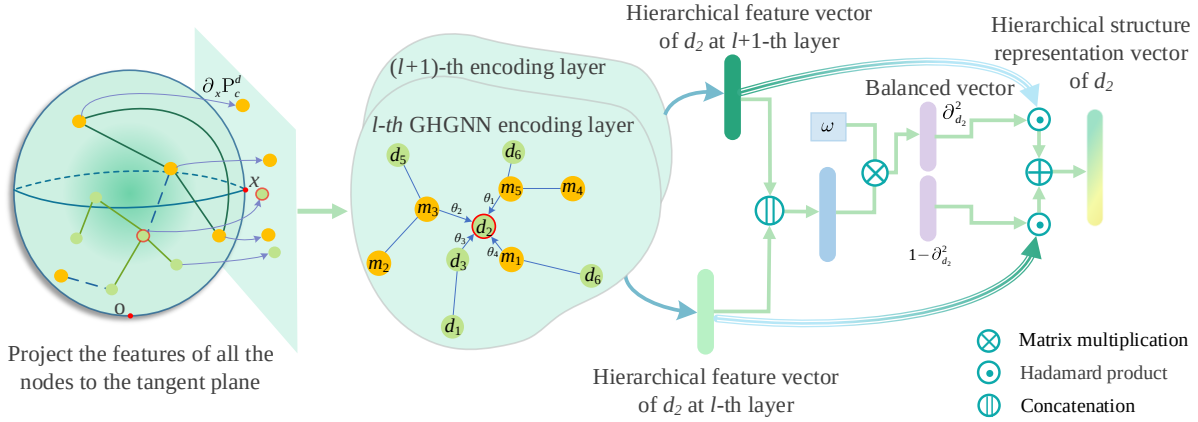


Figure 2: Illustration of node hierarchical feature learning by the designed gating-enhanced hyperbolic graph neural network.

in the number of layers in the graph neural network may lead to the problem of over-smoothing of node features,[37] we design a feature-level gating mechanism to fuse the hierarchical structure features of each node in the previous layer and the current layer. A balanced vector ∂_i^l is constructed as follows:

$$\partial_i^l = \sigma(\omega^l(Z_{feat}^{i,(l)} \,||\, Z_{feat}^{i,(l-1)}) + b^l) \tag{13}$$

where ω^l and b^l are the gating matrix and bias of l-layer, respectively, and $\sigma(\cdot)$ is the sigmoid function. ∂_i^l is used to integrate the information from l-layer and $(l-1)$-layer, and the hierarchical feature of node u_i at l-layer is

$$\tilde{H}_{feat}^{i,(l)} = \partial_i^l \odot Z_{feat}^{i,(l)} + (1 - \partial_i^l) \odot \tilde{H}_{feat}^{i,(l-1)} \tag{14}$$

where $\odot$ represents the operation of Hadamard product. A hierarchical structure features of all drug and microbe nodes is obtained at level L_{numb}, which is renamed Z_{feat}.

Multi-scale topology learning in Euclidean space

Hyperbolic space is good at capturing the hierarchical structure of heterogeneous networks formed by microbes and drugs, while Euclidean space can focus on neighbor topology learning of nodes in heterogeneous networks.

Multi-scale neighbor topology formulation by random walks

A contextual relationship is present between the target node and its neighboring nodes across multiple ranges. To capture this, we construct neighbor topology matrices of varying scales using a neighborhood random walks approach (Fig.3). The neighborhood random walks originates from all drug (microbe) nodes within the network, then moves to adjacent nodes with a defined probability. The microbe-drug neighbor topology matrix $F_k \in R^{N_a \times N_a}$ of scale k is defined as:

$$F_k = \sum_{n=0}^{k} b_{prob}(1 - b_{prob})^n (D^{-1}S)^n \tag{15}$$

where $b_{prob} \in (0, 1)$ is the restart probability, indicating the likelihood of returning to the initial node during the walk. D is the degree matrix calculated from the microbe-drug adjacency matrix S, and k is the length of the random walk. Each row i of F_k records the probability that node u_i reaches all microbe (drug) nodes, with these probabilities reflecting the degree of association between node u_i and these nodes. Therefore, F_k can be interpreted as an adjacency matrix at scale k.

Encoding multi-scale topologies of nodes by graph convolutional networks

Multi-scale neighbor topology plays a crucial role in predicting microbe-drug associations. For k neighbor topologies at different scales , we formulate a multi-scale embedding of all

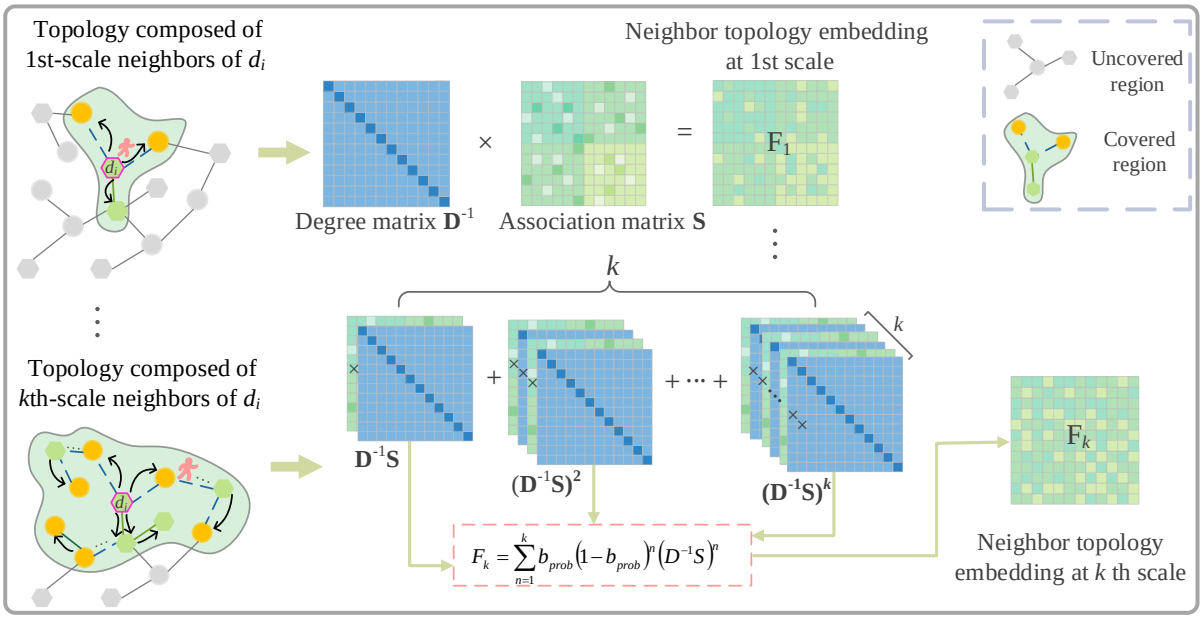


Figure 3: Construction of the multi-scale neighbor topologies by neighborhood random walks on the microbe-drug heterogeneous network.

drug (microbe) nodes based on multi-layer graph convolutional network modules. Given the i-scale neighbor topology , the l-th layer nodes of the graph convolution are embedded as:

$$Z_i^{(l)} = ReLU(F_i Z_i^{(l-1)} W^{(l)}), 2 \leq l \leq N_{layer} \tag{16}$$

where $W^{(l)}$ is the weight matrix of layer l, $Z_i^{(l-1)}$ represents the node feature matrix obtained by graph convolution at layer $l-1$, and N_{layer} is the total number of layers in the graph convolution. Since each graph convolution layer can obtain the node feature matrix under that layer, we concatenate the matrices obtained by graph convolution of N_{layer} layers to obtain $\hat{Z}_i$

$$\hat{Z}_i = concat(Z_i^{(1)}, \ldots, Z_i^{(l)}) \tag{17}$$

Finally, k embeddings of different scales $\hat{Z}_1, \ldots, \hat{Z}_i, \ldots, \hat{Z}_k$ are obtained. These embeddings are fused through 1×1 convolution to form a multi-scale topological features of microbe nodes and drug nodes,

$$Z_{scal} = Conv_{1\times1}(\hat{Z}_1, \ldots, \hat{Z}_i, \ldots, \hat{Z}_k) \tag{18}$$

Fusion of represention from multiple spaces

Z_{feat} and Z_{scal} are hyperbolic property features obtained in hyperbolic space and multi-scale topological features obtained in Euclidean space, respectively. Z_{feat} in hyperbolic space and Z_{scal} in Euclidean space contain unique information specific to each space. Therefore, we design a fusion strategy for multi-space representation. The feature matrix of all microbe and drug nodes, denoted as, X_{feat} contains more detailed characteristics of these nodes. Thus, we concatenate Z_{feat}, Z_{scal}, X_{feat} and to supplement these details and establish MLP layers for weighted fusion. Finally, the fusion attribute representation Z_{fusi} is obtained,

$$Z_{fusi} = \sigma(W^f[Z_{feat} \,||\, Z_{scal} \,||\, X_{feat}] + b^f) \tag{19}$$

where W^f and b^f are the weight matrix and bias, respectively.

Final integration and optimization

We concatenate the attributes Z^i_{fusi} and Z^j_{fusi} of the drug $d_i \in U^{drug}$ and microbe $m_j \in U^{micr}$ into the attribute representation z_{ij} of the node pair d_i-m_j. z_{ij} is then fed into a fully connected layer followed by a softmax layer to obtain the associated probability distribution of d_i and m_j,

$$y_p = softmax(W^p z_{ij} + b^p) \tag{20}$$

where W^p is a weight matrix, and b^p is the bias vector. $y_p[0]$ and $y_p[1]$ represent the probabilities that a drug and a microbe are related and not related, respectively.

During the training process, we employ the Adam optimizer and back propagation to optimize the model. The loss function is defined as the cross-entropy between the true and predicted distributions of the microbe-drug association,

$$loss = -\sum_{i=1}^{N}\sum_{j=1}^{c} t_j \log(y_p[j]) \tag{21}$$

where N is the number of training samples in a batch, $c = 2$. t_j is the actual label value indicating whether the microbe and the drug are associated. If the association between d_i and m_j is known, $t_j = 1$; otherwise, $t_j = 0$.

Experimental evaluations and discussions

Parameter settings

MFLP was developed based on the PyTorch framework and utilized an Nvidia GeForce RTX 3080 graphics card with 10 GB of graphic memory. The training parameters were configured as follows: 120 epochs with a learning rate of 0.0001, and an early stopping strategy was implemented to stop training after 60 epochs. A dropout rate of 0.5 was applied to reduce the effects of overfitting. For the learning hierarchical structure embedding module, the feature dimension of the node was set to 64. In the multi-scale topology features module for microbe and drug nodes, two graph convolutional encoding layers were used, with output feature dimensions of 128 and 64 for the two coding layers, respectively. The random walk restart probability b_{prob} was set to 0.9. The step number of random walks was chosen from {1, 2, 3, 4}, and our model achieved the highest AUC and AUPR when it is 2. The prediction results of the model with different steps were listed in the supplementary table ST2. Additionally, the number of layers of the multi-layer perceptron used to predict the probability of microbe-drug association was set to 2.

Evaluation metrics

We used five-fold cross-validation to evaluate the performance of MFLP in comparison to other methods. The 2470 known microbe-drug associations were all positive samples and were randomly divided into 5 groups of equal size. For each fold, four sets were used for training, while the remaining set was used for testing. For cases where a microbe was not observed to be associated with a drug, the node pair was considered a negative example. We randomly selected an equal number of negative examples as positive examples for training, and the remaining negative examples were reserved for testing.

Evaluation was performed using the area under the receiver operating characteristic (ROC) curve (AUC) and the area under the precision-recall (PR) curve (AUPR) as performance indicators.[38,39] We calculated averaged AUC and AUPR values across the five-fold to assess the overall model performance. Additionally, since top microbe candidates are often selected for further verification by biologists, recall rates were also used as evaluation metrics. Specifically, the recall rate of the top $k \in [30, 60, \ldots, 240]$ was calculated, where a higher recall rate indicates that the prediction model successfully retrieves more actual

microbe-drug associations in the top ranks.

Comparison with other methods

We compared the MFLP model with several state-of-the-artmicrobe-drug association prediction methods, including EGATMDA,[25] GACNNMDA,[40] GCNMDA,[24] GSAMDA,[26] SCSMDA[41] and NGMDA.[28] During the cross-validation process, these six comparison methods were trained and validated using the same training and testing dataset as MFLP. The ROC and PR metrics for all methods are summarized in Table 1. In terms of AUPR, MFLP achieved the highest value (AUPR = 0.767), surpassing NGMDA and EGATMDA by 3.9% and 46.0%, respectively. Compared to GACNNMDA and GCNMDA, MFLP achieved 57.1% and 45.2% improvement, outperformed GSAMDA by 52.0%, and SCSMDA by 42.7%. Additionally, the AUC of MFLP reached the highest value of 0.946, which was 0.2%, 0.6%, 10.3%, 4.3%, 4.4%, and 3.0% higher than the other methods, respectively.

Table 1: AUCs and AUPRs of MFLP and the compared prediction methods.

	EGATMDA	GACNNMDA	GCNMDA	GSAMDA	SCSMDA	NGMDA	Our Method
Average AUC	0.940	0.843	0.903	0.902	0.916	0.944	**0.946**
Average AUPR	0.307	0.196	0.315	0.247	0.340	0.728	**0.767**

For our method and other methods, we performed Wilcoxon test on 1373 AUC pairs and AUPR pairs. The results, as shown in Table 2, indicate that the p-values are less than 0.05, signifying that the AUC and AUPR of MFLP are significantly higher than those of the other comparison methods.

Table 2: The paired Wilcoxon test results for MFLP and each compared method.

	p-value of AUC	p-value of AUPR
SCSMDA	4.8e-158	5.7e-220
GSAMDA	2.0e-157	7.2e-188
GACNNMDA	2.9e-165	9.3e-197
EGATMDA	1.2e-60	1.3e-157
GCNMDA	1.1e-161	1.4e-191
NGMDA	5.0e-10	3.5e-20

Our method (MFLP) exhibits the best performance in terms of both AUC and AUPR, with NGMDA achieving the second-best performance and EGATMDA receiving slightly worse performance. NGMDA enhances node embedded by supervised the encoder and EGATMDA learns node representations using node-level attention to aggregate neighbor node information. Our method aggregates neighbor node information in hyperbolic

space, resulting in MFLP achieving the highest AUC and AUPR. Despite SCSMDA and GCNMDA being based on graph convolutional network, they performed lower than the aforementioned methods. Our approach achieves better performance by integrating multi-scale neighbor topology through graph convolution and neighborhood random walks. GSAMDA achieved slightly worse performance. GACNNMDA, based on convolutional neural networks and graph attention networks, exhibited the worst performance. These methods ignore the node features learned in hyperbolic space, whereas our model effectively integrates the node features from both hyperbolic space and Euclidean space.

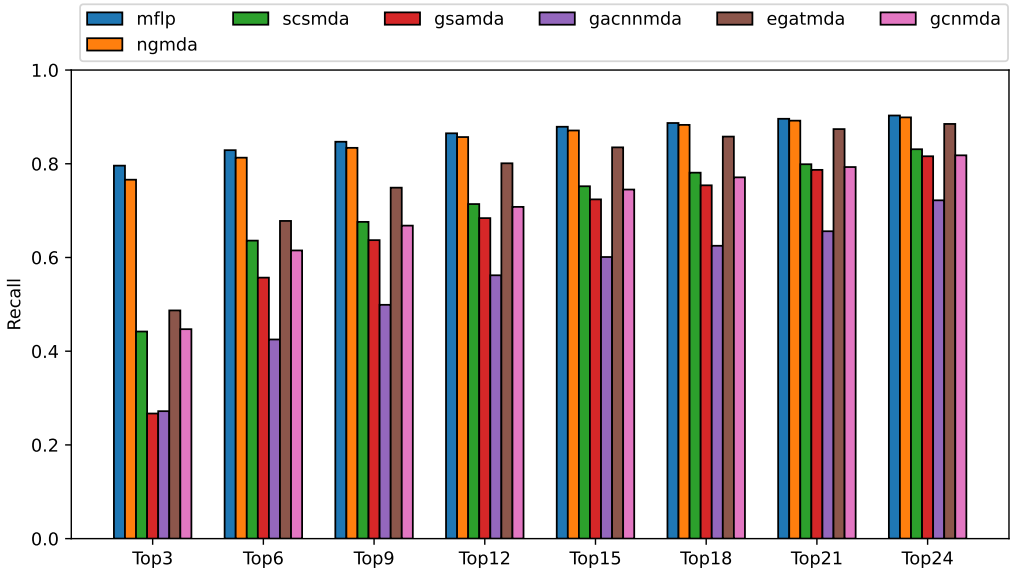


Figure 4: Recall rates of our method and the compared methods for top k ranked candidates.

A higher recall rate for the top k candidate microbe-drug associations signifies that truer microbe-drug associations are included in the candidates. Our method shows better recall rates compared to other comparison methods across different k values (Fig.4). Specifically, when k was set to 30, 60, and 120, MFLP consistently maintained the highest recall rates, reaching 79.6%, 82.9%, and 86.5%, respectively. The second-best performer, NGMDA, achieved recall rates of 76.6%, 81.3%, and 85.7%. EGATMDA ranked third, with recall rates of 48.7%, 67.9%, and 80.1% . In most cases of top k, SCSMDA exhibited slightly higher recall rates than GCNMDA. Specifically, the recall rate of the former was 44.2%, 63.7% and 71.4%, while the latter was 44.7%, 61.5% and 70.8%. GSAMDA achieved lower recall rates, which were 26.7%, 55.8%, and 68.4%, respectively. GACNNMDA consistently achieved the lowest recall rates, with values of 27.2%, 42.5%, and

56.3%.

Ablation experiments

We conducted ablation experiments on the main innovative contributions of our method, including hyperbolic spatial hierarchical structure embedding (HHSE), a gating-enhanced hyperbolic graph neural network (GHGNN), establishment of neighbor topology through random walks (NTRW), and multi-scale neighbor topology learning in Euclidean space (MSTL). As shown in Table 3, the MFLP method containing HHSE, GHGNN, NTRW, and MSTL achieved the best performance, with an AUC of 0.946 and AUPR of 0.767. Removing HHSE resulted in a 1.3% decrease in AUC and an 8.7% decrease in AUPR compared to MFLP, indicating that learning hierarchical embeddings with node categories in hyperbolic space improves prediction performance. Not including GHGNN reduced AUC by 2.3% and AUPR by 14.8%, highlighting the importance of obtaining neighbor information of microbe (drug) nodes on the sphere by learning hierarchical structure features. In Euclidean space, excluding NTRW led to a 4.6% decrease in AUPR of the model, demonstrating that NTRW effectively embeds microbe and drug neighbor nodes of different ranges of topology. Without MSTL, the AUC of MFLP decreased by 0.9% and AUPR by 5.5%. MFLP without GHGNN exhibited more performance degradation than without MSTL, suggesting that learning the attributes of hyperbolic space plays a more critical role. We built a prediction model in which the node feature learning in hyperbolic space (HHSE and GHGNN) was removed. Compared with the complete prediction model, its AUC decreased by 3.2% and AUPR decreased by 14.1%. After the node feature learning in Euclidean space (NTRW and MSTL) have been eliminated, the prediction model's AUC and AUPR decreased by 2.2% and 7.2%, respectively. It indicated the node feature learning in both hyperbolic space and Euclidean space is necessary for the improved prediction performance.

Case studies on ciprofloxacin, moxifloxacin, and amoxicillin

To further validate the capability of MFLP in discovering potential microbe-drug associations, we conducted case studies on moxifloxacin, ciprofloxacin, and amoxicillin. We ranked the microbe candidates based on their predicted microbe-drug association scores in descending order. The top 20 candidate microbes for the three drugs are listed in

Table 3: Ablation experimental results of our method.

HHSE	GHGNN	NTRW	MSTL	Average AUC	Average AUPR
✗	✓	✓	✓	0.933	0.680
✓	✗	✓	✓	0.923	0.619
✓	✓	✗	✓	0.937	0.721
✓	✓	✓	✗	0.937	0.712
✓	✓	✗	✗	0.933	0.695
✗	✗	✓	✓	0.923	0.626
✓	✓	✓	✓	**0.946**	**0.767**

tables 4 to 6.

We used relevant databases and published literature to confirm predicted associations between microbes and drugs. The MDAD provides valuable insights into microbe-drug associations, supported by clinical and experimental studies. In addition, ABiofilm compiles comprehensive information about anti-biofilm drugs through extensive literature research. As shown in Table 4, MDAD lists 3 microbes as top-ranked candidates for the drug moxifloxacin, and ABiofilm also contains 2 microbes. This indicates a genuine association between these microbes and the drug. Furthermore, literature supports 14 of the candidates. Moxifloxacin has been proven to have inhibitory effects on Gram-positive bacteria such as *Staphylococcus aureus*, *Streptococcus sanguis*, and *pneumococcus*.[42–44] Additionally, Gram-negative bacteria like *Pseudomonas aeruginosa* and *Escherichia coli* exhibit varying degrees of sensitivity and resistance to moxifloxacin.[45,46]

Table 4: The top 20 Moxifloxacin-related microbe candidates

Rank	Microbe name	Evidence	Rank	Microbe name	Evidence
1	Candida albicans	MDAD; aBiofilm	11	Salmonella enterica	PMID:22151215
2	Streptococcus mutans	PMID:29160117	12	Vibrio harveyi	Unconfirmed
3	Staphylococcus aureus	PMID:31689174	13	Streptococcus sanguinis	PMID:10629010
4	Escherichia coli	PMID:12714807	14	Streptococcus pneumoniae	PMID:31542319
5	Actinomyces oris	PMID:26538502	15	Haemophilus influenzae	MDAD
6	Pseudomonas aeruginosa	PMID:23662986	16	Acinetobacter baumannii	PMID:12951327
7	Bacillus subtilis	PMID:30036828	17	Stenotrophomonas maltophilia	MDAD; aBiofilm
8	Staphylococcus epidermidis	PMID:31516359	18	Klebsiella pneumoniae	PMID: 33406110
9	Listeria monocytogenes	PMID:28739228	19	Aeromonas hydrophila	PMID: 12821471
10	Streptococcus sanguis	PMID:10629010	20	Burkholderia multivorans	Unconfirmed

The top 20 candidates related to ciprofloxacin are listed in Table 5. MDAD and ABiofilm each have five and three candidates, respectively, indicating the influence of ciprofloxacin on these microbes. Additionally, literature supports 13 of the candidates. Table 6 presents the top 20 candidates for amoxicillin. Two of them appear in MDAD, and ABiofilm verifies two microbe candidates. Literature searches reveal associations

Table 5: The top 20 Ciprofloxacin-related microbe candidates

Rank	Microbe name	Evidence	Rank	Microbe name	Evidence
1	Candida albicans	PMID:31471074	11	Staphylococcus epidermidis	PMID:28481197
2	Pseudomonas aeruginosa	PMID: 33875431	12	Staphylococcus epidermis	PMID:10632381
3	Escherichia coli	MDAD; aBiofilm	13	Burkholderia cenocepacia	PMID:27799222
4	Salmonella enterica	PMID:26933017	14	Plasmodium falciparum	PMID:17214980
5	Streptococcus pneumoniae	MDAD; aBiofilm	15	Listeria monocytogenes	PMID:28355096
6	Staphylococcus aureus	MDAD; aBiofilm	16	Vibrio cholerae	PMID: 21029554
7	Bacillus subtilis	MDAD	17	Burkholderia pseudomallei	PMID:24502667
8	Vibrio harveyi	PMID:27247095	18	Haemophilus influenzae	MDAD
9	Enterococcus faecalis	PMID:27790716	19	Clostridium perfringens	PMID:29978055
10	Streptococcus mutans	PMID:30468214	20	Proteus mirabilis	PMID: 25192738

of 13 microbes with amoxicillin. Among all the predictions, three candidates are not confirmed by the observed evidences among the total 60 candidates and are labeled as "unconfirmed." In summary, the case studies demonstrate that MFLP has a powerful ability to discover potential microbe-drug associations.

Table 6: The top 20 Amoxicillin-related microbe candidates

Rank	Microbe name	Evidence	Rank	Microbe name	Evidence
1	Staphylococcus aureus	PMID: 31396174	11	Enterococcus faecalis	PMID: 8540703
2	Escherichia coli	PMID: 9778163	12	Salmonella enterica	PMID:33446146
3	Candida albicans	PMID: 33806815	13	Klebsiella pneumoniae	PMID: 29489447
4	Pseudomonas aeruginosa	PMID: 18499192	14	Staphylococcus epidermis	Unconfirmed
5	Streptococcus mutans	PMID: 33657016	15	Clostridium perfringens	MDAD
6	Bacillus subtilis	aBiofilm	16	Micrococcus luteus	PMID: 22589499
7	Staphylococcus epidermidis	PMID: 27491399	17	Actinomyces oris	PMID: 34684101
8	Vibrio harveyi	PMID: 32056705	18	Aeromonas hydrophila	PMID: 2387294
9	Listeria monocytogenes	aBiofilm; MDAD	19	Haemophilus influenzae	PMID: 15215115
10	Streptococcus pneumoniae	PMID: 35748540	20	Vibrio cholerae	PMID: 37616193

Prediction of novel microbes related to drugs

We used all known microbe-drug associations, as well as an equal number of selected negative examples, to train the MFLP model. Afterward, The model was applied to 1373 drugs in order to predict new microbe-drug associations. To assist biologists in identifying reliable drug-associated microbes, we included the top 20 ranked microbe candidates in Supplementary Table ST1.

Conclusions

We proposed a novel method that improved the performance of predicting drug-related candidate microbes, and the method encoded and integrated the hierarchical structure and multi-scale neighbor topologies in multiple spaces. In hyperbolic space, the constructed hierarchical embeddings are beneficial for encoding the hierarchical structures composed

of microbe and drug nodes. A hyperbolic heterogeneous neural network with a gating mechanism aggregated the features of both microbe and drug neighbors for the target node. The feature-level gate is proposed to assign higher weight to the more informative features of each node of the current graph neural encoding layer and integrate the node hierarchical features of the current encoding layer and those of the previous layer. In terms of the heterogeneous network in Euclidean space, the neighbor topological embeddings with multi scales are formed by adjusting the ranges that a random walker walks in the network. Multiple graph convolutional networks are constructed for the multi-scale neighbor topologies to encode their specific topologies and attributes respectively. The 5-fold cross-validation results indicated that MFLP achieved higher AUC and AUPR than 5 advanced prediction methods. MFLP also demonstrated its superior for the top ranked candidates, which means more actual microbe-drug associations appeared in these candidates. The case studies on three drugs, including Ciprofloxacin, Moxifloxacin and Amoxicillin, demonstrated MFLP is able to discover the potential candidate microbes for drugs.

Data and Software Availability

The source codes and datasets are freely available at https://github.com/pingxuan-hlju/MFLP.

Acknowledgement

This work was supported by the Natural Science Foundation of Heilongjiang Province [LH2023F044]; Natural Science Foundation of China [62372282, 62172143]; Natural Science Foundation of Guangdong Province [2024A1515010176]; STU Scientific Research Initiation Grant [NTF22032].

Supporting Information Available

The authors declare no competing financial interest.

- The top 20 ranked microbe candidates related to drugs are recorded in Supplemental Table ST1
- The prediction results of the model with different steps of random walks were listed in the supplementary table ST2.

References

(1) Maraz, K. M.; Khan, R. A. An overview on impact and application of microorganisms on human health, medicine and environment. *GSC Biological and Pharmaceutical Sciences* **2021**, *16*, 089–104.

(2) Wilson, M.; Wilson, P. J.; Wilson, M.; Wilson, P. J. Microbes and infectious diseases. *Close Encounters of the Microbial Kind: Everything You Need to Know About Common Infections* **2021**, 3–48.

(3) Pathak, K.; Saikia, R.; Gogoi, U.; Das, A. Potential application of microbes for human health and welfare. *World J. Pharm. Pharm. Sci* **2020**, *10*, 514–524.

(4) Kitamoto, S.; Nagao-Kitamoto, H.; Hein, R.; Schmidt, T.; Kamada, N. The bacterial connection between the oral cavity and the gut diseases. *Journal of dental research* **2020**, *99*, 1021–1029.

(5) Peng, X.; Cheng, L.; You, Y.; Tang, C.; Ren, B.; Li, Y.; Xu, X.; Zhou, X. Oral microbiota in human systematic diseases. *International journal of oral science* **2022**, *14*, 14.

(6) Willis, J. R.; Gabaldón, T. The human oral microbiome in health and disease: from sequences to ecosystems. *Microorganisms* **2020**, *8*, 308.

(7) Abe, F. C.; Kodaira, K.; Motta, C. d. C. B.; Barberato-Filho, S.; Silva, M. T.; Guimarães, C. C.; Martins, C. C.; Lopes, L. C. Antimicrobial resistance of microorganisms present in periodontal diseases: A systematic review and meta-analysis. *Frontiers in Microbiology* **2022**, *13*, 961986.

(8) Mira, A.; Simon-Soro, A.; Curtis, M. Role of microbial communities in the pathogenesis of periodontal diseases and caries. *Journal of clinical periodontology* **2017**, *44*, S23–S38.

(9) Yahaya, T.; Obaroh, I.; Oldele, E. The roles of microorganisms in the pathogenesis and prevalence of Diabetes: a review. *diabetes* **2017**, *1*, 3.

(10) Castrillon, C. A.; Hincapie, J. P.; Yepes, F. L.; Roldan, N.; Moreno, S. M.; Contreras, A.; Botero, J. E. Occurrence of red complex microorganisms and A ggregatibacter actinomycetemcomitans in patients with diabetes. *Journal of investigative and clinical dentistry* **2015**, *6*, 25–31.

(11) Aldars-Garcia, L.; Chaparro, M.; Gisbert, J. P. Systematic review: the gut microbiome and its potential clinical application in inflammatory bowel disease. *Microorganisms* **2021**, *9*, 977.

(12) Liu, S.; Zhao, W.; Lan, P.; Mou, X. The microbiome in inflammatory bowel diseases: from pathogenesis to therapy. *Protein & cell* **2021**, *12*, 331–345.

(13) Zhang, Y.; Xiang, X.; Zhou, S.; Dindar, D. A.; Wood, S.; Zhang, Z.; Shan, B.; Zhao, L. Relationship between pathogenic microorganisms and the occurrence of esophageal carcinoma based on pathological type: a narrative review. *Expert Review of Gastroenterology & Hepatology* **2023**, *17*, 353–361.

(14) Yano, Y.; Etemadi, A.; Abnet, C. C. Microbiome and cancers of the esophagus: a review. *Microorganisms* **2021**, *9*, 1764.

(15) Yahya, E. B.; Abdulsamad, M. A.; Allaq, A. A. Recent advances in the role of microorganisms in cancer incidence: Mechanisms and health precautions. *Biomedical Research and Therapy* **2021**, *8*, 4525–4539.

(16) Uddin, T. M.; Chakraborty, A. J.; Khusro, A.; Zidan, B. R. M.; Mitra, S.; Emran, T. B.; Dhama, K.; Ripon, M. K. H.; Gajdács, M.; Sahibzada, M. U. K.; others Antibiotic resistance in microbes: History, mechanisms, therapeutic strategies and future prospects. *Journal of infection and public health* **2021**, *14*, 1750–1766.

(17) Chinemerem Nwobodo, D.; Ugwu, M. C.; Oliseloke Anie, C.; Al-Ouqaili, M. T.; Chinedu Ikem, J.; Victor Chigozie, U.; Saki, M. Antibiotic resistance: The challenges and some emerging strategies for tackling a global menace. *Journal of clinical laboratory analysis* **2022**, *36*, e24655.

(18) Jian, Z.; Zeng, L.; Xu, T.; Sun, S.; Yan, S.; Yang, L.; Huang, Y.; Jia, J.; Dou, T.

Antibiotic resistance genes in bacteria: Occurrence, spread, and control. *Journal of basic microbiology* **2021**, *61*, 1049–1070.

(19) Rather, M. A.; Gupta, K.; Mandal, M. Microbial biofilm: formation, architecture, antibiotic resistance, and control strategies. *Brazilian Journal of Microbiology* **2021**, 1–18.

(20) Bungau, S.; Tit, D. M.; Behl, T.; Aleya, L.; Zaha, D. C. Aspects of excessive antibiotic consumption and environmental influences correlated with the occurrence of resistance to antimicrobial agents. *Current Opinion in Environmental Science & Health* **2021**, *19*, 100224.

(21) Zhu, L.; Duan, G.; Yan, C.; Wang, J. Prediction of microbe-drug associations based on chemical structures and the KATZ measure. *Current Bioinformatics* **2021**, *16*, 807–819.

(22) Long, Y.; Luo, J. Association mining to identify microbe drug interactions based on heterogeneous network embedding representation. *IEEE journal of biomedical and health informatics* **2020**, *25*, 266–275.

(23) Zhu, L.; Wang, J.; Li, G.; Hu, X.; Ge, B.; Zhang, B. Predicting microbe-drug association based on similarity and semi-supervised learning. *Am J Biochem Biotechnol* **2021**, *17*, 50–8.

(24) Long, Y.; Wu, M.; Kwoh, C. K.; Luo, J.; Li, X. Predicting human microbe–drug associations via graph convolutional network with conditional random field. *Bioinformatics* **2020**, *36*, 4918–4927.

(25) Long, Y.; Wu, M.; Liu, Y.; Kwoh, C. K.; Luo, J.; Li, X. Ensembling graph attention networks for human microbe–drug association prediction. *Bioinformatics* **2020**, *36*, i779–i786.

(26) Tan, Y.; Zou, J.; Kuang, L.; Wang, X.; Zeng, B.; Zhang, Z.; Wang, L. GSAMDA: a computational model for predicting potential microbe–drug associations based on graph attention network and sparse autoencoder. *BMC bioinformatics* **2022**, *23*, 492.

(27) Deng, L.; Huang, Y.; Liu, X.; Liu, H. Graph2MDA: a multi-modal variational graph embedding model for predicting microbe–drug associations. *Bioinformatics* **2022**, *38*, 1118–1125.

(28) Xuan, P.; Gu, J.; Cui, H.; Wang, S.; Toshiya, N.; Liu, C.; Zhang, T. Multi-scale topology and position feature learning and relationship-aware graph reasoning for prediction of drug-related microbes. *Bioinformatics* **2024**, *40*, btae025.

(29) Sun, Y.-Z.; Zhang, D.-H.; Cai, S.-B.; Ming, Z.; Li, J.-Q.; Chen, X. MDAD: a special resource for microbe-drug associations. *Frontiers in cellular and infection microbiology* **2018**, *8*, 424.

(30) Hattori, M.; Tanaka, N.; Kanehisa, M.; Goto, S. SIMCOMP/SUBCOMP: chemical structure search servers for network analyses. *Nucleic acids research* **2010**, *38*, W652–W656.

(31) Alrasheed, H.; Dragan, F. F. Core–periphery models for graphs based on their δ-hyperbolicity: An example using biological networks. *Journal of Algorithms & Computational Technology* **2017**, *11*, 40–57.

(32) Adcock, A. B.; Sullivan, B. D.; Mahoney, M. W. Tree-Like Structure in Large Social and Information Networks. 2013 IEEE 13th International Conference on Data Mining. 2013; pp 1–10.

(33) Clauset, A.; Moore, C.; Newman, M. E. Hierarchical structure and the prediction of missing links in networks. *Nature* **2008**, *453*, 98–101.

(34) Ravasz, E.; Barabási, A.-L. Hierarchical organization in complex networks. *Physical review E* **2003**, *67*, 026112.

(35) Weber, S.; Zöngür, B.; Araslanov, N.; Cremers, D. Flattening the Parent Bias: Hierarchical Semantic Segmentation in the Poincaré Ball. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024; pp 28223–28232.

(36) Clevert, D.; Unterthiner, T.; Hochreiter, S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). 4th International Conference on Learning

Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings. 2016.

(37) Wang, X.; Zhu, M.; Bo, D.; Cui, P.; Shi, C.; Pei, J. AM-GCN: Adaptive multi-channel graph convolutional networks. Proceedings of the 26th ACM SIGKDD International conference on knowledge discovery & data mining. 2020; pp 1243–1253.

(38) Yang, S.; Berdine, G. The receiver operating characteristic (ROC) curve. *The Southwest Respiratory and Critical Care Chronicles* **2017**, *5*, 34–36.

(39) Flach, P.; Kull, M. Precision-recall-gain curves: PR analysis done right. *Advances in neural information processing systems* **2015**, *28*.

(40) Ma, Q.; Tan, Y.; Wang, L. GACNNMDA: a computational model for predicting potential human microbe-drug associations based on graph attention network and CNN-based classifier. *BMC bioinformatics* **2023**, *24*, 35.

(41) Tian, Z.; Yu, Y.; Fang, H.; Xie, W.; Guo, M. Predicting microbe–drug associations with structure-enhanced contrastive learning and self-paced negative sampling strategy. *Briefings in Bioinformatics* **2023**, *24*, bbac634.

(42) Dubois, J.; Dubois, M. Levonadifloxacin (WCK 771) exerts potent intracellular activity against Staphylococcus aureus in THP-1 monocytes at clinically relevant concentrations. *Journal of medical microbiology* **2019**, *68*, 1716–1722.

(43) Hoogkamp-Korstanje, J.; Roelofs-Willemse, J. Comparative in vitro activity of moxifloxacin against Gram-positive clinical isolates. *Journal of Antimicrobial Chemotherapy* **2000**, *45*, 31–39.

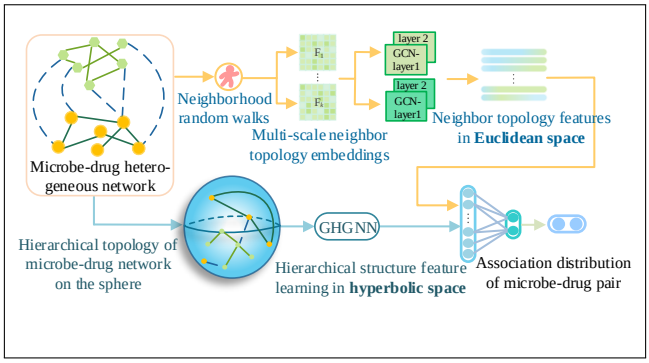
(44) Dalhoff, A.; Bowker, K.; MacGowan, A. Comparative evaluation of eight in vitro pharmacodynamic models of infection: activity of moxifloxacin against Escherichia coli and Streptococcus pneumoniae as an exemplary example. *International journal of antimicrobial agents* **2020**, *55*, 105809.

(45) Oldenburg, C. E.; Lalitha, P.; Srinivasan, M.; Rajaraman, R.; Ravindran, M.; Mascarenhas, J.; Borkar, D. S.; Ray, K. J.; Zegans, M. E.; McLeod, S. D.; others Emerg-

ing moxifloxacin resistance in Pseudomonas aeruginosa keratitis isolates in South India. *Ophthalmic epidemiology* **2013**, *20*, 155–158.

(46) Lemmen, S.; Häfner, H.; Klik, S.; Lütticken, R.; Zolldann, D. Comparison of the bactericidal activity of moxifloxacin and levofloxacin against Staphylococcus aureus, Staphylococcus epidermidis, Escherichia coli and Klebsiella pneumoniae. *Chemotherapy* **2003**, *49*, 33–35.

TOC Graphic


Microbe-drug hetero-geneous network
Neighborhood random walks
Multi-scale neighbor topology embeddings
layer 2 GCN-layer1
layer 2 GCN-layer1
Neighbor topology features in Euclidean space
Hierarchical topology of microbe-drug network on the sphere
GHGNN
Hierarchical structure feature learning in hyperbolic space
Association distribution of microbe-drug pair