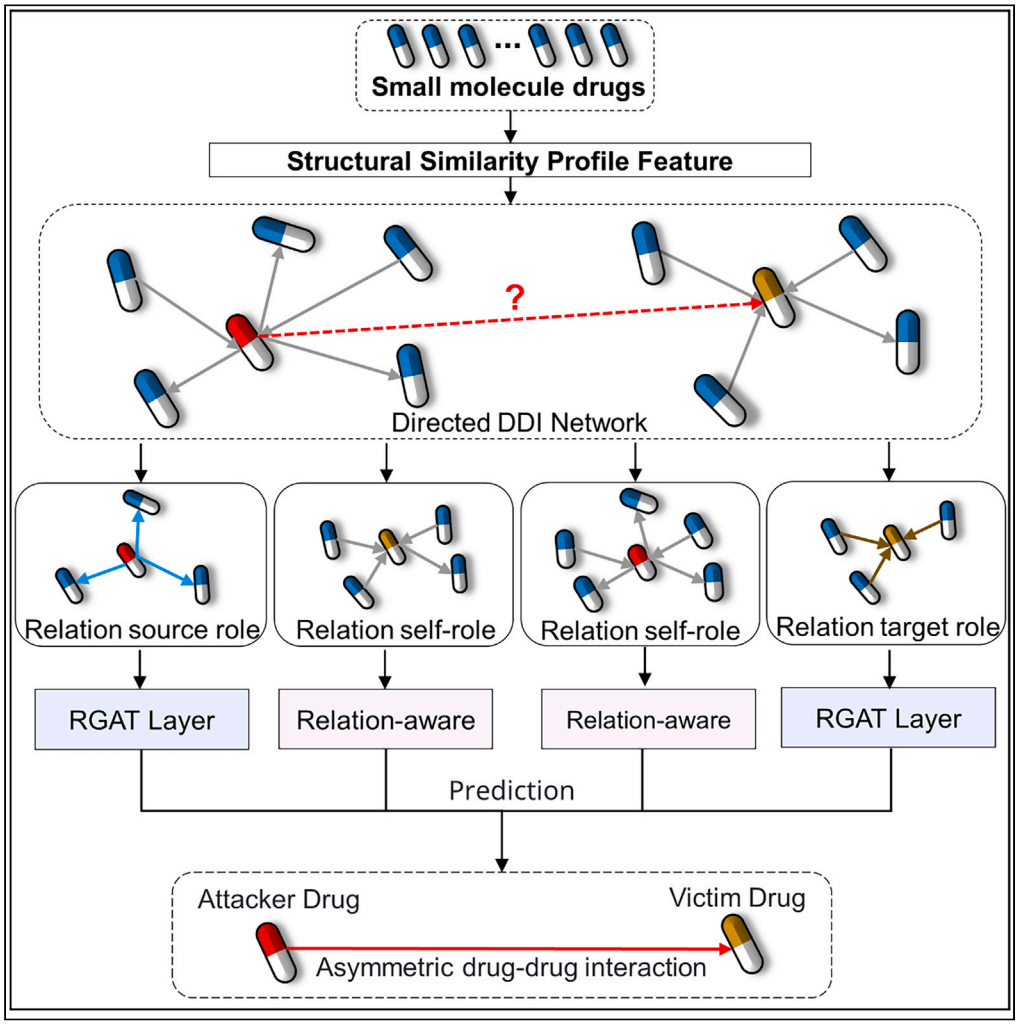
iScience

CellPress
OPEN ACCESS

Article

DRGATAN: Directed relation graph attention aware network for asymmetric drug-drug interaction prediction



Dehai Zhang, Zhengwu Wang, Di Zhao, Jin Li

lijin@ynu.edu.cn

Highlights

Directed graph of asymmetric drug interactions

Invasiveness and vulnerability information of drugs in asymmetric DDI

Different relations role embedding of drugs in directed DDI graph

The impact of different interactions on asymmetric DDI information

Zhang et al., iScience 27, 109943
June 21, 2024
https://doi.org/10.1016/j.isci.2024.109943

iScience

CellPress
OPEN ACCESS

Article

DRGATAN: Directed relation graph attention aware network for asymmetric drug-drug interaction prediction

Dehai Zhang,[1] Zhengwu Wang,[1] Di Zhao,[1] and Jin Li[1,2,*]

SUMMARY

In scenarios involving the treatment of complex or coexisting diseases with multiple drugs, the potential for severe adverse drug reactions in patients necessitates the identification of potential drug-drug interactions (DDIs). Most existing computational methods have not taken into account the asymmetry and relation types of drug interactions caused by the relation information between drugs, which may lead to missing information in embedded learning. Therefore, this paper proposes a directed relation graph attention aware network (DRGATAN) to predict asymmetric drug interactions. DRGATAN leverages an encoder to learn multi-relational role embeddings of drugs across different types of relations. The experimental results show that DRGATAN's performance is superior to recognized advanced methods. The visualization demonstrates the effect of utilizing asymmetric information, and the case analysis validates the reliability of the proposed method. This study provides guidance for predicting asymmetric drug interactions.

INTRODUCTION

With the increase of complex diseases and the improvement of patient drug resistance, multi-drug combination therapy has become a mainstream treatment method.[1,2] While multi-drug therapy offers considerable therapeutic promise, the concomitant administration of multiple drugs introduces the potential for drug-drug interactions (DDIs), wherein the activity of individual drugs may be altered, leading to potentiation or attenuation of their effects.[3] These unforeseen DDIs can precipitate serious adverse drug reactions (ADEs) in patients,[4,5] culminating in severe morbidity or mortality.[6,7] Conversely, beneficial interactions can synergistically enhance treatment efficacy. Consequently, effective prediction of potential adverse DDIs before drug co-administration is imperative. However, traditional biological or pharmacological methods for detecting DDIs are laborious and resource-intensive.[8,9]

In recent years, researchers have proposed various techniques, among which the use of deep neural networks or graph neural networks to predict DDIs has received widespread attention.[10] For example, DeepDDI[11] uses deep neural networks to predict interactions, and MR-GNN[12] uses multiple graph convolutional layers to learn drug features. META-DDIE[13] mines interpretable drug substructures by employing a neural network-based encoder-decoder framework to learn drug representations. This approach enhances performance in predicting rare DDIs while also providing interpretability. DDIMDL[14] adopts a joint deep neural network framework to learn cross-modality representations of drug-drug pairs for predicting DDI events. DANN-DDI[15] utilizes graph embedding techniques and deep attention neural networks to learn representations of drug pairs from multiple drug feature networks. It then employs deep neural networks to accurately predict potential DDIs. Furthermore, graph-based learning methods have witnessed rapid development, reshaping DDI prediction into a link prediction task on DDI networks. In general, the process of learning entity embeddings from networks often involves knowledge graph embedding methods, which leverage various types of relations to generate embeddings of entities, including multi-relational embeddings. For instance, Tiresias[16] harnesses TransH[17] and HolE[18] to embed drugs and their relations. The approach using graph neural networks learns entity embeddings by aggregating information from neighboring nodes. For example, GCNMK[19] uses graph convolutional networks (GCN)[20] to learn the potential topological neighborhood information of entities in DDI networks with two different types of relations. MRCGNN[21] hierarchically integrates drug molecular structure information with DDI event graph interaction information. Leveraging relational GCN for multi-relational contrastive learning, it captures implicit event information, thereby enhancing performance in predicting rare DDIs. In contrast, knowledge graph embedding methods[22] focus on directly learning entity embeddings across multiple relations while disregarding entity neighborhood information.

In contrast to traditional methods, deep learning-based computational approaches offer the capability of conducting large-scale preliminary DDI screening with advantages such as reduced time, lower costs, and enhanced accuracy.[23,24] Through a comprehensive survey of existing literature[25,26] and referencing relevant literature reviews, we categorize advanced computational methods into two primary groups: those based on drug molecular structural features and those based on network structures.

[1]The Key Laboratory of Software Engineering of Yunnan Province, School of Software, Yunnan University, Kunming 650091, P.R. China
[2]Lead contact
*Correspondence: lijin@ynu.edu.cn
https://doi.org/10.1016/j.isci.2024.109943

Previous methods were primarily based on drug molecular structural features,[27] utilizing structural similarity, and interaction spectral fingerprint similarity to predict potential DDIs. For instance, the HNAI framework[28] incorporated four similarity features to compute drug-drug pair similarity, while the INDI (inferring drug interactions) approach[29] utilized seven distinct drug-drug similarity measures for this purpose. Zhang et al.[30] creatively integrated multi-source data and drug pair similarities to predict potential DDIs. Their approach involved incorporating nine types of drug-related data and employing three representative methods to construct diverse prediction models utilizing varied information, ultimately achieving high-precision performance. Subsequently, researchers shifted toward employing graph neural networks to learn drug molecular structural features. Asada et al.[31] utilized GCN[32] to encode molecular structures, thereby enhancing DDI extraction from textual data. GCN-BMP[33] (graph convolutional network with bond-aware message propagation) learns specific features and abstract features through GCN and attention mechanisms to obtain paired drug representations. Additionally, MHCADDI[34] solely relied on the side effect type and molecular structure of drugs to predict adverse interactions.

Methods based on network structure primarily center on the topology of the network. These methods organize DDI data into a DDI network, where nodes represent drugs and edges symbolize different types of interactions. Zitnik et al.[35] leveraged knowledge graph embedding technology and applied RESCAL[36] to infer the most suitable relations between drugs in multi-relational data. KGNN[37](knowledge graph neural network) utilized a comprehensive biomedical knowledge graph (KG), encompassing enzymes, targets, and genes, to learn drug features. Decagon[38] learned associations between drug pairs and side effects by constructing multimodal graphs to glean drug features. DDI-MDAE[39] achieved unified representation of drugs by simultaneously learning multiple drug feature networks. GoGNN[40] employed the interaction graph of molecular graphs to extract structured entity features of drug molecules. Although these advanced methods have achieved good results, they directly utilize graph neural networks or embedding methods to learn drug entity embeddings, ignoring the inequality caused by asymmetric interaction relations between drugs and the influence of different interaction relations on the propagation of neighborhood information among drug entities. This leads to incomplete information in embedding learning.

In general, existing methods tend to treat interacting drug pairs as having pharmacological equivalence, overlooking the pharmacological asymmetry of drug interactions. Many biological experiments have demonstrated the pharmacological asymmetry of drug interactions. The study by Wicha et al.[41] elucidated that a significant proportion of drug combinations involving antifungal and non-antifungal agents manifest unidirectional interactions. For example, terbinafine can antagonize the action of amphotericin B via the ergosterol pathway in a unidirectional manner. Asymmetric interactions can also influence the sequence of drug administration in specific scenarios. For example, administering vincristine before cyclophosphamide has been shown to enhance anti-tumor activity, while concurrent administration does not exhibit additive effects.[42] Clinical trials involving cisplatin and 5-fluorouracil have demonstrated that different drug administration sequences yield varied effects and risks.[43] Similar strategies have been validated across various treatments, highlighting the critical importance of determining the sequence of drug administration in multi-drug therapies. Consequently, it is imperative to develop new deep learning models capable of predicting potential asymmetric DDIs. The insights garnered from research on asymmetric interactions can significantly enhance our understanding of the mechanisms underlying drug interactions and offer rational strategies for optimizing multi-drug therapies. The DGAT-DDI method[44] pioneered the prediction of asymmetric DDIs by constructing a directed DDI network and employing graph attention networks (GAT)[45] to learn asymmetric information regarding drug interactions. Similarly, MAVGAE[46] leverages multimodal data and variational graph autoencoder for predicting asymmetric DDIs. However, both approaches overlook the critical importance of modeling multi-relational information for capturing asymmetric information effectively.

To address the aforementioned issues, we have devised a method that comprehensively incorporates multi-relational information, asymmetric information, and network structure information for DDI prediction. In this paper, we propose a new framework called Directed Relation Graph Attention Aware Network, which is noted as DRGATAN (see Figure 1). Our approach involves constructing a directed DDI network based on asymmetric relations. We then aim to learn the distinct role characteristics between drugs under asymmetric relations, while simultaneously capturing the multi-relational network structure information of drugs by propagating neighbor information of drug entities across different relations.

RESULTS

Dataset

From the DrugBank 5.1.9 version, we collected DDI data on asymmetric interactions, focusing primarily on small molecule drugs approved by the FDA (Food and Drug Administration). We have designated the collected dataset of asymmetric interactions as Asymmetric-DDI. We filtered out small molecule drugs for which we couldn't obtain molecular fingerprints from the original DDI data, as we required molecular Morgan fingerprints to calculate the structural similarity profile (SSP) between drugs as the initial feature vector of drug nodes. As a result, we gathered data on 1,876 drugs and 218,917 asymmetric interactions, spanning across 95 types of relations.

Experimental setup

The encoder of our model, DRGATAN, comprises RGAT layers and a relation-aware network. These components are utilized to capture the source (target) role of relations and the self-role representation of drugs, respectively. We employ $l = 2$ RGAT layers to individually learn the embeddings of the relation source (target). Each layer is composed of $k = 16$ heads for multi-head attention. We initially input the initial features $\boldsymbol{x} \in \mathbb{R}^{100}$ that have undergone dimensionality reduction through principal-component analysis (PCA), resulting in a feature vector dimension of 100. Each asymmetric interaction $r \in R$ is represented by a learnable matrix $\mathbf{W}^r \in \mathbb{R}^{100\times16}$. The RGAT layer's final output is a hidden feature vector of dimension 16. In the relation self-role embedding, the input for the relation-aware network also consists of the initial feature vector $\boldsymbol{x} \in \mathbb{R}^{100}$. The network produces an output feature vector with a dimension of 16.

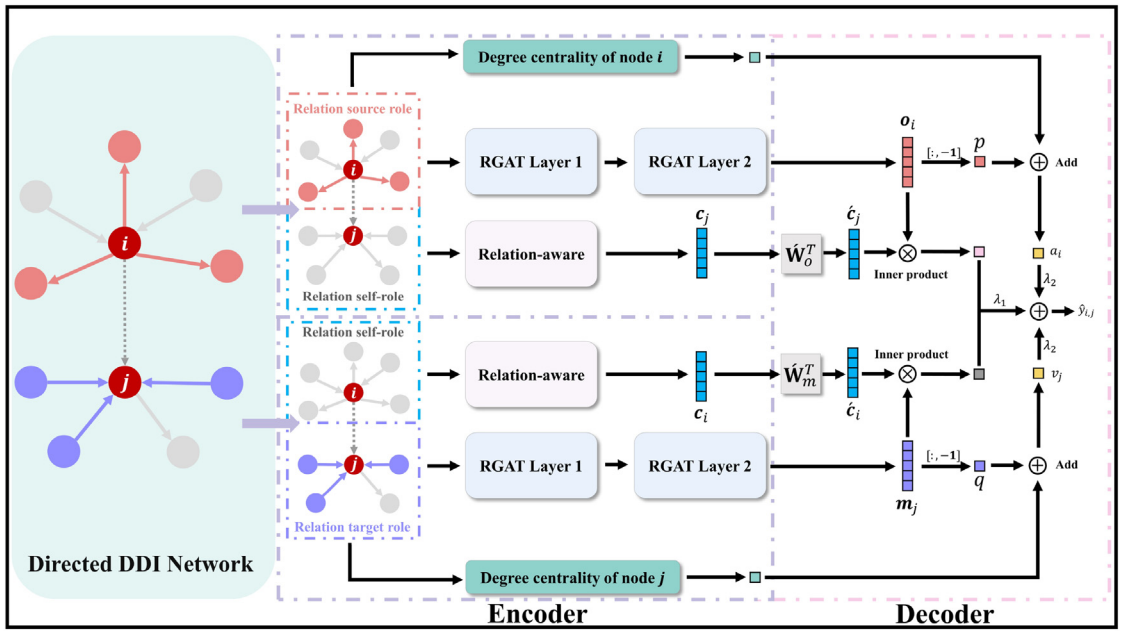


Figure 1. DRGATAN model architecture

The overall structure of the model consists of an encoder (light purple dashed box) and a decoder (light pink dashed box). The relation source role embedding is obtained by aggregating the outgoing nodes of drug *i* (three light pink nodes). The relation target role embedding is obtained by aggregating the incoming nodes of drug *j* (three purple nodes). The relation self-role embedding of drug *i* is obtained by aggregating the features of neighboring nodes (gray nodes) of drug *j* through a relation-aware network.

During training, parameters are adjusted to achieve optimal performance. We set the dropout rate to 0.6 during the embedding process. The learning rate *lr* is 1*e*–2. Each training session comprises 200 epochs, and the model is trained using the Adam optimizer.[47] For the hyper-parameters λ_1 and λ_2, they are set to $\lambda_1 = 0.4$ and $\lambda_2 = 0.1$. λ_1 controls the similarity of relation role embedding between constrained nodes. λ_2 controls the aggressiveness and vulnerability of constrained nodes under different relation role embeddings. λ_1 and λ_2 jointly coordinate the relation source role and invasiveness, as well as the relation target role and vulnerability.

Baseline model

The following baseline models are utilized for comparison and performance evaluation alongside our method.

DeepDDI[11]: it uses the SSP of drugs as input to a deep neural network (DNN) for predicting the type of interaction between drugs.

DPDDI[48]: extracting network structure characteristics of drugs from the DDI network through GCN, and subsequently utilizing deep neural networks for binary DDI prediction.

GCNMK[19]: using GCN to learn the features of drugs for binary DDI prediction by learning DDI graphs composed of "increase" and "decrease" interactions separately.

GoGNN[40]: using GNN and dual attention mechanism to learn the features of structured entity graph and entity interaction graph for DDI prediction.

GCN-BMP[33]: using GCN and attention mechanism to learn the low-dimensional features of drug chemical structures, constructing a neural network based on HoLE[18] for DDI prediction.

DGATDDI[44]: use directed DDI graph construction and GAT to learn the asymmetric interaction information between drug nodes to achieve specific direction DDI prediction.

In this study, we used 5-fold cross-validation to train the model. During each fold process, 1-fold is allocated as the test set. Subsequently, 80% of the remaining 4-folds are assigned to the training set, while the remaining 20% serves as the validation set. All experiments were conducted on our Asymmetric-DDI dataset. Negative samples are generated through random sampling, encompassing instances where asymmetric interactions do not exist between drugs and cases where the direction of asymmetric interactions is incorrect. To ensure balance in the training data, the number of positive and negative samples is kept equal at all stages.

Experimental results

In this section, we analyze the performance of the proposed method by comparing the results with the baseline. Table 1 shows the average and standard deviation (SD) of the six indicators obtained by 5-fold cross-validation on the Asymmetric-DDI dataset, including AUROC, AUPRC, ACC, F1, precision, and recall. The best indicator results are highlighted in bold. From this table, we observe that DRGATAN exhibits

Table 1. Performance comparison in binary DDI prediction (%)

Model	AUROC	AUPRC	ACC	F1	PRECISION	RECALL
DeepDDI[11]	91.90±0.15	87.55±0.29	86.79±0.11	86.95±0.12	85.73±0.40	88.21±0.51
DPDDI[48]	91.64±0.05	89.96±0.08	84.41±0.07	85.07±0.05	81.61±0.22	88.84±0.27
GCNMK[19]	95.73±0.08	95.10±0.06	90.18±0.17	90.23±0.13	89.54±0.44	90.93±0.26
GoGNN[40]	94.93±0.07	94.06±0.14	89.30±0.17	89.45±0.13	88.31±1.04	90.65±1.20
GCN-BMP[33]	88.57±0.60	84.93±0.55	81.58±0.74	82.25±0.88	79.19±1.01	85.61±2.30
DGATDDI[44]	96.24±0.01	95.47±0.01	90.50±0.04	90.63±0.05	89.43±0.13	91.87±0.23
OURs	**98.58±0.02**	**98.30±0.02**	**95.02±0.04**	**95.05±0.04**	**94.53±0.16**	**95.57±0.21**

Values correspond to the mean (standard deviation). The best results appear in bold.

significantly better performance. Specifically, compared to other methods, DRGATAN has at least a 2.34% improvement in AUROC, 2.83% improvement in AUPRC, 4.52% improvement in ACC, and 4.42% improvement in F1. These results underscore the effectiveness of our method.

We evaluate the performance of the proposed method by comparing it with baseline methods using the results obtained from DRGATAN. Compared with the second-best result of DGATDDI, our method has an ACC of 95.02%, which is an absolute increase of 4.52%. The reason is that DGATDDI did not consider the information on relation types in DDI. In contrast, DRGATA not only learns the asymmetric interaction information between drugs but also integrates asymmetric relation features into drug embedding. DPDDI and GCNMK are also network structure-based methods. Compared with them, DRGATAN also achieved better results. For example, it improved by 10.61% in ACC and 9.98% in F1 compared with DPDDI. While GCNMK considers two different interaction types of DDI graphs, its embedding process remains incomplete. Additionally, the embedding of DPDDI and GCNMK relies on network structure, underscoring the effectiveness of our method in learning useful network structure information from a directed DDI network.

Furthermore, the methods of DeepDDI, GoGNN, and GCNBMP all incorporate drug molecular features. DRGATAN also outperformed these methods. Compared with DeepDDI, our performance may have been enhanced by incorporating network structure features and asymmetric interaction information. Both GoGNN and GCNBMP utilize molecular structures to learn drug features. It is worth noting that GoGNN also integrates network structure and edge features to augment molecular structure features. Compared with it, ACC has improved by 5.72%. This demonstrates that DRGATAN is capable of effectively learning network structure information and asymmetric interaction information from the directed DDI network. In summary, DRGATAN can extract rich information from the DDI network and exhibits excellent performance in predicting non-symmetric interaction DDI.

Analysis of drug features

To investigate the impact of different initial drug features on model performance, six common drug molecular fingerprint features were extracted to analyze the model's performance. The comparison results with different drug features are shown in Table 2, where the best result is highlighted in bold.

The drug molecular fingerprints include MACCS, atom pair (atom pair fingerprint), ECFP4, ECFP6, RDKit (RDKit Fingerprint), and topological (topological torsion fingerprint). The number of bits for extracting different types of molecular fingerprints varies to prevent the influence of different feature dimensions on experimental results. For fairness, we adopt the same procedure as SSP and apply PCA to reduce the feature dimensions of molecular fingerprints to 100. From the table, it can be observed that SSP features achieved the best performance compared to the other six types of drug molecular fingerprints. The model improved by at least 0.44% on AUROC, at least 0.55% on AUPRC, at least 0.9% on ACC, at least 0.92% on F1, at least 0.77% on PRECISION, and at least 1.27% on RECALL. We have the following

Table 2. The impact of different drug features on model performance (%)

Features	AUROC	AUPRC	ACC	F1	PRECISION	RECALL
MACCS	98.47 ± 0.01	98.16 ± 0.01	94.80 ± 0.01	94.82 ± 0.01	94.45 ± 0.04	95.20 ± 0.05
Atom Pair	98.15 ± 0.01	97.79 ± 0.02	94.12 ± 0.04	94.13 ± 0.05	93.97 ± 0.19	94.30 ± 0.29
ECFP4	98.26 ± 0.01	97.91 ± 0.02	94.37 ± 0.04	94.39 ± 0.05	94.02 ± 0.21	94.76 ± 0.29
ECFP6	98.24 ± 0.02	97.87 ± 0.02	94.34 ± 0.02	94.37 ± 0.03	93.92 ± 0.08	94.82 ± 0.10
RDKit	98.33 ± 0.02	98.01 ± 0.03	94.55 ± 0.05	94.57 ± 0.05	94.18 ± 0.20	94.97 ± 0.24
Topological	98.14 ± 0.01	97.75 ± 0.01	94.17 ± 0.05	94.20 ± 0.06	93.76 ± 0.08	94.64 ± 0.18
SSP	**98.58±0.02**	**98.30±0.02**	**95.02±0.04**	**95.05±0.04**	**94.53±0.16**	**95.57±0.21**

Values correspond to the mean (standard deviation). The best results appear in bold.

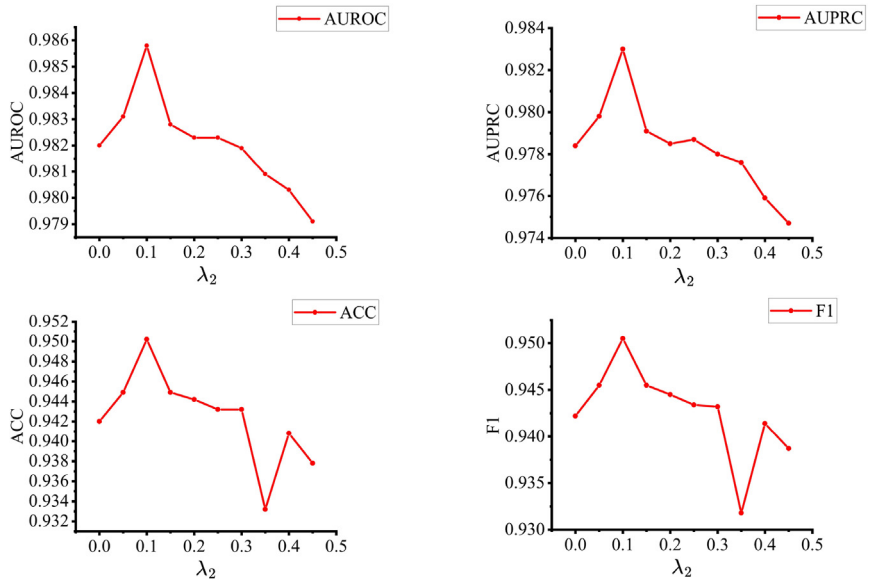


Figure 2. The effect of hyperparameter λ_2

Analyze the influence of adjusting the λ_2 parameter on the equilibrium between the relation source role embedding and invasiveness, alongside the relation target role embedding and vulnerability, on model performance.

observations: (1) atom pair and topological molecular fingerprints performed the worst compared to other drug features. This could be due to atom pair being limited by the quality of the fingerprints influenced by the selection of atom pairs in the structure, and Topological may not accurately describe certain structural features. On the other hand, SSP can provide a more detailed structural description reflecting crucial structural features of drug molecules. (2) The performance of models using ECFP6, ECFP4, and RDKit fingerprints increases in that order, but they may all be influenced by the path length parameter affecting fingerprint quality. Additionally, ECFP may not provide a rich enough description of molecular ring structures, hence resulting in lower performance compared to SSP. (3) The MACCS fingerprint feature achieved suboptimal results. This could be because MACCS only considers 166 predefined structural keys and one invalid bit. For simple structure descriptions like those of small molecule drugs, this concise representation is often sufficient to capture key structural features. PCA reduces its dimensionality with less loss of feature information compared to other fingerprints. However, it may not be as flexible as SSP and may not adapt well to variations in features across different molecules, resulting in slightly lower performance. In conclusion, the experimental results indicate that using SSP to generate drug features yields better performance in predicting asymmetric DDIs, demonstrating superior representational capabilities that enhance predictive accuracy.

Parameter analysis

We studied the effect of hyperparameter λ_2 on performance in Equation 7. The objective is to determine the optimal balance parameters for the relationship source role embedding and invasiveness, as well as the relationship target role embedding and vulnerability, to optimize the model's performance. In the model configuration, we defined $\lambda_1 + \lambda_2 = 0.5$. Both jointly balance the relation source role embedding with invasiveness and the relation target role embedding with vulnerability. This parameter λ_2 is used to constrain the invasiveness and vulnerability of different relation role embeddings. Specifically, the model training settings remain consistent, all other parameters are fixed, only the value of λ_2 is changed, $\lambda_2 \in [0, 0.45]$, and the change interval is 0.05. The results are shown in Figure 2. The figure illustrates that as the value of λ_2 increases, all four indicators initially exhibit an upward trend, with the peak occurring at $\lambda_2 = 0.1$. Subsequently, as the value of λ_2 is further increased, the four indicators exhibit a downward trend. When the performance is optimal, $\lambda_2 = 0.1$ is a relatively small value, which is expected. We need to constrain the interference caused by nodes with excessively high centrality degrees on asymmetric information. The direction of asymmetric interactions is jointly determined by the relation role embedding, invasiveness, and vulnerability in order to achieve higher accuracy. Excessively high levels of invasiveness and vulnerability can lead to the dominance of degree centrality in asymmetric DDI prediction, resulting in a weakening of the multi-relational role embedding effect and a subsequent decline in overall performance.

Ablation study

Studying the effect of asymmetric interaction relations on the relation role between drug nodes with asymmetric information, as well as the invasiveness and vulnerability of nodes, is an integral component of our method. Therefore, ablation experiments are used to study their contribution to the prediction results. We have implemented three methods to remove key modules to verify their effectiveness. The

Table 3. Ablation experiment results (%)

Model	AUROC	AUPRC	ACC	F1
(w/o)DS	98.21	97.82	94.37	94.38
(w/o)RA	97.84	97.30	93.56	93.61
(w/o)RF	96.26	95.63	90.45	90.62
All	98.58	98.30	95.02	95.05

(w/o)DS represents the method without considering degree centrality. (w/o)RA represents the method without considering the relation-aware network structure information of nodes. (w/o)RF represents the method without considering the feature information of asymmetrical relations.

experimental settings here are consistent with those that achieve the best performance. The analysis is based on the indicators of AUROC, AUPRC, ACC, and F1. Overall, the complete method outperformed all of its ablation models. The experimental results are shown in Table 3.

In order to assess the importance of asymmetric relations on the embedding of role information in asymmetric information, we exclude relation features (w/o) RF from consideration for asymmetric DDI prediction. We can see from the table that the complete method has an absolute increase of 4.57% on ACC and 4.43% on F1, and its impact on the prediction results is the greatest. This result shows that the contribution of asymmetric interaction relations to the difference in roles caused by relations and the contribution of asymmetric information is significant in the embedding of relation source (target) roles. In order to investigate the impact of relation-aware network structure information, we excluded the relation-aware network component (w/o) RA. Compared with it, the complete method has an increase of 1.46% on ACC and 1.44% on F1. This further underscores the effectiveness of network topology structural features in modeling asymmetric information, thereby enhancing the performance of the model. To study whether the degree of nodes is invasiveness or vulnerability, we removed this component (w/o) DS. It can be seen that the complete method has improved performance compared to it. F1 increased by 0.67%, ACC increased by 0.65%, and AUPRC increased by 0.48%. This reflects the importance of capturing how the number of neighbors under different roles of drugs affects the degree of asymmetric interaction, all of which contribute to the influence of asymmetric information. The results suggest that intrusiveness and fragility play a significant role in predicting asymmetric interactions in DDI.

Case study

Relation role embeddings visualization

To better understand the modeling of asymmetric information, we analyze the visualization of drug relation role embeddings and compare and discuss the effects of attention mechanisms and relation information in modeling. Specifically, the asymmetry of the asymmetric interaction $i \rightarrow j$ between drug i and drug j is jointly characterized by the relation source role of i and the relation target role of j. The relation source role embedding $\boldsymbol{o}_i$ of drug i (attacker) and the relation self role embedding $\boldsymbol{c}_j$ of drug j are close in the embedding space. The relation target role embedding $\boldsymbol{m}_j$ of drug j (victim) and the relation self role embedding $\boldsymbol{c}_i$ of drug i are close in the embedding space. As an example, taking *Dabrafenib* → *Delavirdine*, we employ t-SNE[49] to visualize the relation role embeddings, aiming to analyze the impact of attention mechanisms and relation information. To alleviate the impact of imbalance in the number of outgoing and incoming nodes for drugs, we aim to select drug pairs where the number of outgoing nodes for the attacker drug is close to the number of incoming nodes for the victim drug. Dabrafenib, as the attacker drug, has 616 outgoing interactions. Delavirdine, as the victim drug, has 628 incoming interactions.

Figure 3 displays the visualization of relation role embeddings for the drug pair *Dabrafenib* → *Delavirdine* under different scenarios. In the embedding space, the attacker dabrafenib's relation source role (depicted as a red right triangle) and its outgoing neighbor's relation self-role (illustrated as green dots), as well as the victim delavirdine's relation target role (represented by a purple left triangle) and its incoming neighbor's relation self-role (depicted as pink dots), are displayed. From Figure 3A, two separate and closely clustered groups can be observed. Dabrafenib and its outgoing neighbors form a closely clustered, while delavirdine and its incoming neighbors also form a closely clustered. Furthermore, dabrafenib and delavirdine are distantly separated from each other. This implies that DRGATAN effectively leverages attention mechanisms and relation information to enhance the quality of relation role embeddings. From Figure 3B, it is evident that when only ignoring relation information, dabrafenib and delavirdine are closely clustered together in space, making it challenging to differentiate between them. Additionally, they are distinctly separated from their respective clusters. Furthermore, their respective clusters disperse into smaller clusters. This indicates that attention mechanisms primarily concentrate on local neighborhoods. Without directional relation information, capturing relation role embeddings with asymmetric information becomes impossible, rendering the modeling of such asymmetry challenging. Figure 3C demonstrates the impact of ignoring attention mechanisms exclusively. Although dabrafenib and delavirdine are situated in distinct clusters, there remains a tendency for them to be in proximity to each other compared to Figure 3A. Ignoring relation information without accounting for the significance of neighborhoods may still lead to the inability to differentiate between relation source role embeddings and relation target role embeddings, particularly in scenarios involving numerous similar interaction relations. Figure 3D also indirectly suggests the potential occurrence of this scenario, as evidenced by two closely situated clusters and the close clustering of dabrafenib and delavirdine in space. Without taking into account attention mechanisms and relation information, relation role embeddings that solely aggregate features of neighboring nodes cannot adequately capture asymmetric information. In short, DRGATAN effectively utilizes

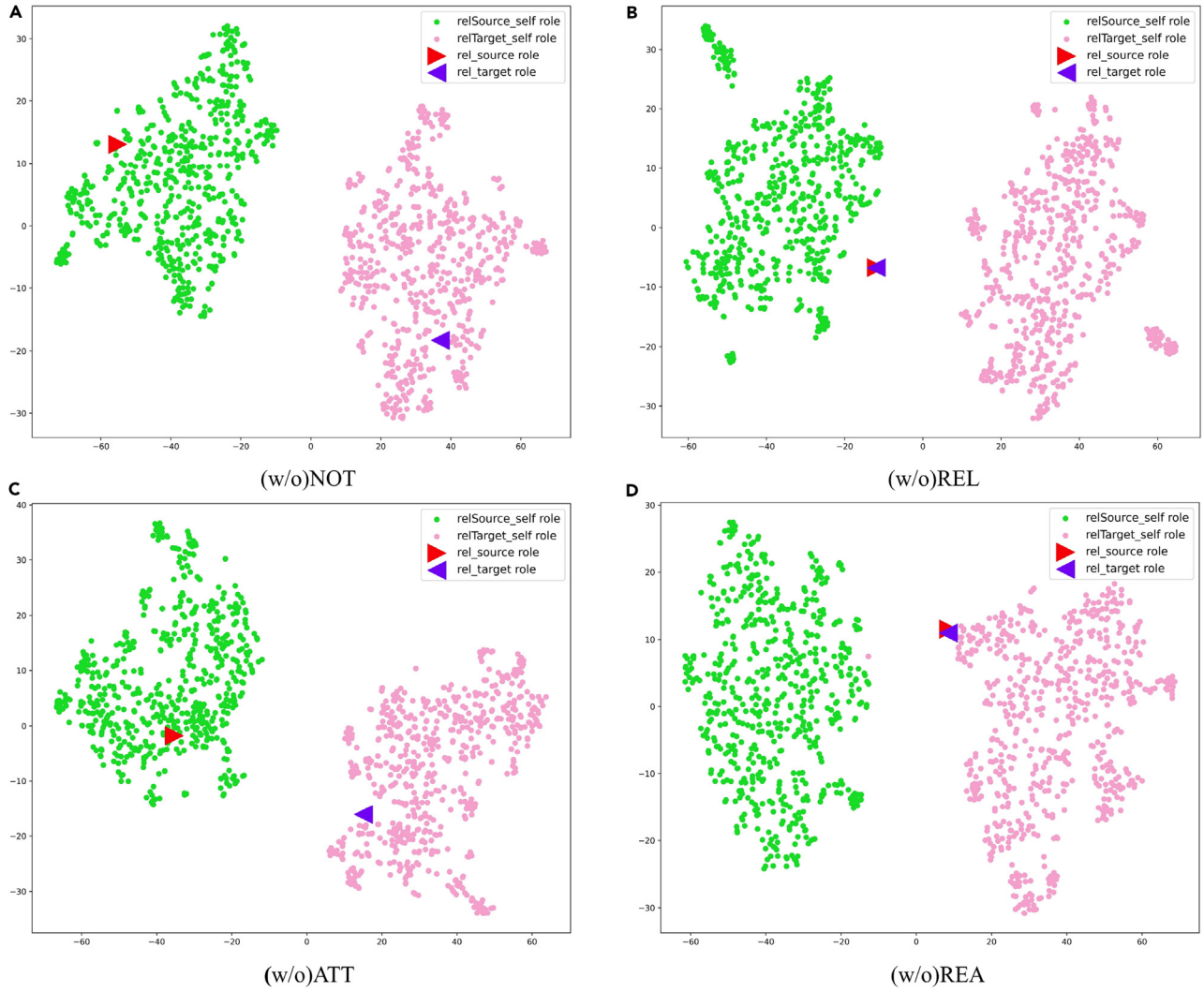


Figure 3. Visualize the relation roles embedding of interacting drug pairs using t-SNE

(A) (w/o)NOT represents the simultaneous consideration of attention mechanism and relation information.

(B) (w/o)REL represents the exclusion of relation information.

(C) (w/o)ATT represents the exclusion of the attention mechanism.

(D) (w/o)REA represents the exclusion of both the attention mechanism and relation information.

attention mechanisms and relation information to learn more representative relation role embeddings, thereby enhancing its ability to capture and represent asymmetric interaction information.

Similarly, using moclobemide as an example, we conducted an analysis of the performance of relation role embeddings when individual drugs act as attackers and victims. When moclobemide acts as the attacker drug, it has 290 outgoing interactions, whereas the victim drug, it has 232 incoming interactions.

Figure 4 displays the relation source role of moclobemide as the attacker rug (depicted as a red right triangle) and the relation self-role of its outgoing neighbors (represented by green dots), as well as the relation target role of moclobemide as the victim drug (depicted as a purple left triangle) and the relation self-role of its incoming neighbors (depicted as pink dots) in the embedding space. When both attention mechanisms and relation information are considered simultaneously, Figure 4A reveals the presence of two completely independent clusters. The relation source role and relation target role of moclobemide are distantly separated from each other, while their corresponding relation self-roles are closely clustered together. This observation further underscores the effectiveness of attention mechanisms and relation information in the analysis. Similarly, when either attention mechanisms or relation information are disregarded, the relation source role of moclobemide tends to closely cluster with the cluster containing its relation target role, thus hindering the effective modeling of asymmetric information. It is worth noting that Figure 4C illustrates that without considering attention mechanisms, the accurate capture of similarity between the relation source role and its corresponding relation self-role is hindered. This further underscores the pivotal role of attention mechanisms in learning

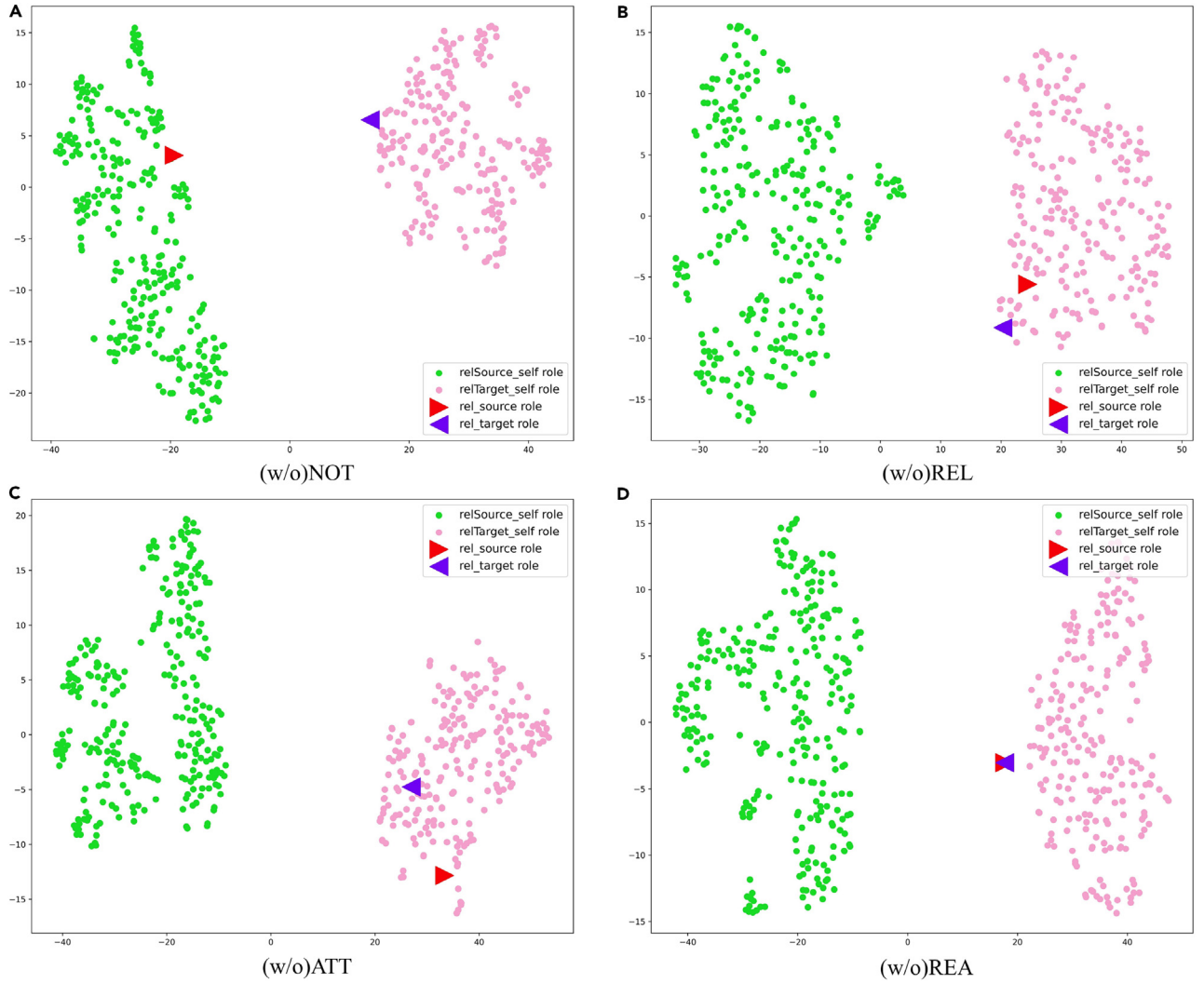


Figure 4. Visualizing the relation roles embedding of individual drugs using t-SNE

(A) (w/o)NOT represents the simultaneous consideration of attention mechanism and relation information.

(B) (w/o)REL represents the exclusion of relation information.

(C) (w/o)ATT represents the exclusion of the attention mechanism.

(D) (w/o)REA represents the exclusion of both the attention mechanism and relation information.

asymmetric information. In conclusion, the visual analysis highlights the critical role of attention mechanisms and relation information in effectively modeling asymmetric information.

Predict new asymmetric DDIs

In this section, we perform a case study to assess the effectiveness of DRGATAN in predicting the actual potential of undetected DDIs. In the current drug dataset, many unmarked drug pairs may have undetected DDIs. We use the current data to train the model based on the transduction setting and predict the interaction between these unobserved drug pairs, which do not exist in the current data. Subsequently, a list of DDIs is generated based on the predicted scores, arranged in descending order. For these unknown drug pairs, a higher predicted score indicates a greater likelihood of interaction. The top 20 DDI data are selected and shown in Table 4, where the drug pairs are affected by drug to drug, and the direction indicates whether the direction of asymmetric interaction between the two drugs is predicted correctly. By verifying the predicted DDIs in the list with the latest version of the DrugBank (5.1.10) database, we found that 16 out of the 20 newly predicted asymmetric interactions can be confirmed. There are 13 completely correct predictions and 3 interaction directions (the second, fourth, and thirteenth) that are predicted incorrectly. However, there are 4 newly predicted DDIs that could not be found in the new database for validation, which may indicate that they have not been confirmed to exist, or that they were predicted incorrectly. According to the case study, our DRGATAN can effectively discover asymmetric interactions between potential drugs.

Table 4. Top 20 asymmetric DDIs

Rank	Drug *i* (attacker)	Drug *j* (victim)	Description	Direction
1	Cephalexin	Ramelteon	Cephalexin may decrease the excretion rate of Ramelteon which could result in a higher serum level.	√
2	Ceftibuten	Quinethazone	Quinethazone may increase the excretion rate of Ceftibuten which could result in a lower serum level and potentially a reduction in efficacy.	×
3	Ouabain	Nedaplatin	N.A.	–
4	Ceftriaxone	Quinethazone	Quinethazone may increase the excretion rate of Ceftriaxone which could result in a lower serum level and potentially a reduction in efficacy.	×
5	Cephalexin	Nedocromil	Cephalexin may decrease the excretion rate of Nedocromil which could result in a higher serum level.	√
6	Thiabendazole	Diazepam	Thiabendazole may decrease the excretion rate of Diazepam which could result in a higher serum level.	√
7	Digitoxin	Efinaconazole	N.A.	–
8	Cephalexin	Megestrol acetate	Cephalexin may decrease the excretion rate of Megestrol acetate which could result in a higher serum level.	√
9	Rifabutin	Chlorambucil	N.A.	–
10	Zanamivir	Diazepam	Zanamivir may decrease the excretion rate of Diazepam which could result in a higher serum level.	√
11	Cephalexin	Metoclopramide	Cephalexin may decrease the excretion rate of Metoclopramide which could result in a higher serum level.	√
12	Nedocromil	Diazepam	Nedocromil may decrease the excretion rate of Diazepam which could result in a higher serum level.	√
13	Diazepam	Trifluridine	Trifluridine may decrease the excretion rate of Diazepam which could result in a higher serum level.	×
14	Ampicillin	Diazepam	Ampicillin may decrease the excretion rate of Diazepam which could result in a higher serum level.	√
15	Cephalexin	Ampicillin	Cephalexin may decrease the excretion rate of Ampicillin which could result in a higher serum level.	√
16	Topotecan	Ropinirole	N.A.	–
17	Pamidronic acid	Clorazepic acid	Pamidronic acid may decrease the excretion rate of Clorazepic acid which could result in a higher serum level.	√
18	Methotrexate	Flurazepam	Methotrexate may decrease the excretion rate of Flurazepam which could result in a higher serum level.	√
19	Flutamide	Diazepam	Flutamide may decrease the excretion rate of Diazepam which could result in a higher serum level.	√
20	Cephalexin	Isotretinoin	Cephalexin may decrease the excretion rate of Isotretinoin which could result in a higher serum level.	√

N.A. indicates that there is no relevant description available for the given DDI at the moment. Direction indicates whether the direction of the asymmetric interaction is predicted correctly.

DISCUSSION

In this paper, we introduce a DRGATAN designed for predicting asymmetric DDIs. DRGATAN effectively learns embeddings representing the roles of drugs as both sources and targets under various relation types. Additionally, it employs a relation-aware network to capture structural embeddings of a multi-relation directed DDI network, thus characterizing relation self-role embeddings. These three types of role embeddings are combined to model asymmetric interactions comprehensively. Moreover, DRGATAN incorporates measures of invasiveness and vulnerability of drug role embeddings based on degree centrality and the number of neighbors. In experimental evaluations, DRGATAN outperforms baseline methods, demonstrating its efficacy in predicting asymmetric DDI tasks.

In the future, we plan to tackle the existing limitations of our approach. Specifically, we recognize that solely considering SSP drug features and low-order neighbors may restrict the modeling of asymmetric information. To overcome this limitation, we intend to incorporate multi-source drug information and consider path information to capture various types of asymmetric interactions among high-order nodes. Furthermore, the imbalance in asymmetric interaction data also needs to be addressed in the future to improve the model prediction accuracy.

Limitations of the study

Although the proposed method performed well in experiments, there are still some limitations that need to be considered. The method only considers the asymmetric information between nodes and their first-order neighboring nodes based on directed relationships when modeling asymmetric information, ignoring the influence of high-order nodes and their directional relationships on asymmetric information. Path information can be introduced to capture the relationship information between high-order nodes and help the model more comprehensively capture asymmetric information. Future research directions can focus on addressing these limitations and exploring more effective methods to improve model performance.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Problem definition
  - Drug structure similarity features
  - Relation role encoding
  - Invasiveness and vulnerability
  - Prediction and loss
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109934.

ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of China (62362066, 62366059) and the Natural Science Foundation of Yunnan Province (202001BB050052).

AUTHOR CONTRIBUTIONS

D. Zhang and J.L. conceived and designed the experiments. Z.W., D. Zhao, and J.L. performed the experiments. Z.W. and D. Zhao conducted the data and analyses. Z.W. wrote the paper. D. Zhang and. J.L. reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 26, 2023
Revised: March 21, 2024
Accepted: May 6, 2024
Published: May 8, 2024

REFERENCES

1. Bansal, M., Yang, J., Karan, C., Menden, M.P., Costello, J.C., Tang, H., Xiao, G., Li, Y., Allen, J., Zhong, R., et al. (2014). A community computational challenge to predict the activity of pairs of compounds. Nat. Biotechnol. *32*, 1213–1222. https://doi.org/10.1038/nbt.3052.
2. Han, K., Jeng, E.E., Hess, G.T., Morgens, D.W., Li, A., and Bassik, M.C. (2017). Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat. Biotechnol. *35*, 463–474. https://doi.org/10.1038/nbt.3834.
3. Prueksaritanont, T., Chu, X., Gibson, C., Cui, D., Yee, K.L., Ballard, J., Cabalu, T., and Hochman, J. (2013). Drug-drug interaction studies: regulatory guidance and an industry perspective. AAPS J. *15*, 629–645. https://doi.org/10.1208/s12248-013-9470-x.
4. Edwards, I.R., and Aronson, J.K. (2000). Adverse drug reactions: definitions, diagnosis, and management. Lancet *356*, 1255–1259. https://doi.org/10.1016/S0140-6736(00) 02799-9.
5. Lazarou, J., Pomeranz, B.H., and Corey, P.N. (1998). Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA *279*, 1200–1205. https://doi.org/10.1001/jama.279.15.1200.
6. Vilar, S., Uriarte, E., Santana, L., Lorberbaum, T., Hripcsak, G., Friedman, C., and Tatonetti, N.P. (2014). Similarity-based modeling in large-scale prediction of drug-drug interactions. Nat. Protoc. *9*, 2147–2163. https://doi.org/10.1038/nprot.2014.151.
7. Shtar, G., Rokach, L., and Shapira, B. (2019). Detecting drug-drug interactions using artificial neural networks and classic graph similarity measures. PLoS One *14*, e0219796. https://doi.org/10.1371/journal.pone.0219796.
8. Whitebread, S., Hamon, J., Bojanic, D., and Urban, L. (2005). Keynote review: *in vitro* safety pharmacology profiling: an essential tool for successful drug development. Drug Discov. Today *10*, 1421–1433. https://doi.org/10.1016/S1359-6446(05)03632-9.
9. Ding, Y., Tang, J., and Guo, F. (2019). Identification of drug-side effect association via multiple information integration with centered kernel alignment. Neurocomputing *325*, 211–224. https://doi.org/10.1016/j.neucom.2018.10.028.
10. Yue, X., Wang, Z., Huang, J., Parthasarathy, S., Moosavinasab, S., Huang, Y., Lin, S.M., Zhang, W., Zhang, P., and Sun, H. (2020). Graph embedding on biomedical networks: methods, applications and evaluations. Bioinformatics *36*, 1241–1251. https://doi.org/10.1093/bioinformatics/btz718.
11. Ryu, J.Y., Kim, H.U., and Lee, S.Y. (2018). Deep learning improves prediction of drug-drug and drug-food interactions. Proc. Natl. Acad. Sci. USA *115*, E4304–E4311. https://doi.org/10.1073/pnas.1803294115.
12. Xu, N., Wang, P., Chen, L., Tao, J., and Zhao, J. (2019). Mr-gnn: Multi-resolution and dual graph neural network for predicting structured entity interactions. Preprint at arXiv. https://doi.org/10.24963/ijcai.2019/551.
13. Deng, Y., Qiu, Y., Xu, X., Liu, S., Zhang, Z., Zhu, S., and Zhang, W. (2022). META-DDIE: predicting drug-drug interaction events with few-shot learning. Briefings Bioinf. *23*, bbab514. https://doi.org/10.1093/bib/bbab514.
14. Deng, Y., Xu, X., Qiu, Y., Xia, J., Zhang, W., and Liu, S. (2020). A multimodal deep learning framework for predicting drug-drug interaction events. Bioinformatics *36*, 4316–4322. https://doi.org/10.1093/bioinformatics/btaa501.
15. Liu, S., Zhang, Y., Cui, Y., Qiu, Y., Deng, Y., Zhang, Z., and Zhang, W. (2023). Enhancing drug-drug interaction prediction using deep attention neural networks. IEEE ACM Trans. Comput. Biol. Bioinf *20*, 976–985. https://doi.org/10.1109/TCBB.2022.3172421.
16. Abdelaziz, I., Fokoue, A., Hassanzadeh, O., Zhang, P., and Sadoghi, M. (2017). Large-scale structural and textual similarity-based mining of knowledge graph to predict drug-drug interactions. J. Web Semant. *44*, 104–117. https://doi.org/10.1016/j.websem.2017.06.002.
17. Wang, Z., Zhang, J., Feng, J., and Chen, Z. (2014). Knowledge graph embedding by translating on hyperplanes. Proc. AAAI Conf. Artif. Intell. *28*. https://doi.org/10.1609/aaai.v28i1.8870.
18. Nickel, M., Rosasco, L., and Poggio, T. (2016). Holographic embeddings of knowledge graphs. Proc. AAAI Conf. Artif. Intell. *30*. https://doi.org/10.1609/aaai.v30i1.10314.
19. Wang, F., Lei, X., Liao, B., and Wu, F.-X. (2022). Predicting drug-drug interactions by graph convolutional network with multi-kernel. Briefings Bioinf. *23*, bbab511. https://doi.org/10.1093/bib/bbab511.
20. Kipf, T.N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. Preprint at arXiv. https://doi.org/10.48550/arXiv.1609.02907.
21. Xiong, Z., Liu, S., Huang, F., Wang, Z., Liu, X., Zhang, Z., and Zhang, W. (2023). Multi-relational contrastive learning graph neural network for drug-drug interaction event prediction. Proc. AAAI Conf. Artif. Intell. *37*, 5339–5347. https://doi.org/10.1609/aaai.v37i4.25665.
22. Wang, Q., Mao, Z., Wang, B., and Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. IEEE Trans. Knowl. Data Eng. *29*, 2724–2743. https://doi.org/10.1109/TKDE.2017.2754499.
23. Zhang, W., Jing, K., Huang, F., Chen, Y., Li, B., Li, J., and Gong, J. (2019). SFLLN: a sparse feature learning ensemble method with linear neighborhood regularization for predicting drug-drug interactions. Inf. Sci. *497*, 189–201. https://doi.org/10.1016/j.ins. 2019.05.017.
24. Zhang, Y.P., and Zou, Q. (2020). PPTPP: a novel therapeutic peptide prediction method using physicochemical property encoding and adaptive feature representation learning. Bioinformatics *36*, 3982–3987. https://doi.org/10.1093/bioinformatics/btaa275.
25. Lin, X., Dai, L., Zhou, Y., Yu, Z.-G., Zhang, W., Shi, J.-Y., Cao, D.-S., Zeng, L., Chen, H., Song, B., et al. (2023). Comprehensive evaluation of deep and graph learning on drug-drug interactions prediction. Briefings Bioinf. *24*, bbad235. https://doi.org/10.1093/bib/bbad235.
26. Qiu, Y., Zhang, Y., Deng, Y., Liu, S., and Zhang, W. (2022). A comprehensive review of computational methods for drug-drug interaction detection. IEEE ACM Trans. Comput. Biol. Bioinf *19*, 1968–1985. https://doi.org/10.1109/TCBB.2021.3081268.
27. Vilar, S., Harpaz, R., Uriarte, E., Santana, L., Rabadan, R., and Friedman, C. (2012). Drug-drug interaction through molecular structure similarity analysis. J. Am. Med. Inf. Assoc. *19*, 1066–1074. https://doi.org/10.1136/amiajnl-2012-000935.
28. Cheng, F., and Zhao, Z. (2014). Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. J. Am. Med. Inf. Assoc. *21*, e278–e286. https://doi.org/10.1136/amiajnl-2013-002512.
29. Gottlieb, A., Stein, G.Y., Oron, Y., Ruppin, E., and Sharan, R. (2012). INDI: a computational framework for inferring drug interactions and their associated recommendations. Mol. Syst. Biol. *8*, 592. https://doi.org/10.1038/msb.2012.26.
30. Zhang, W., Chen, Y., Liu, F., Luo, F., Tian, G., and Li, X. (2017). Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. BMC Bioinf. *18*, 18. https://doi.org/10.1186/s12859-016-1415-9.
31. Asada, M., Miwa, M., and Sasaki, Y. (2018). Enhancing drug-drug interaction extraction from texts by molecular structure information. Preprint at arXiv. https://doi.org/10.48550/arXiv.1805.05593.
32. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P.S. (2021). A comprehensive survey on graph neural networks. IEEE Transact. Neural Networks Learn. Syst. *32*, 4–24. https://doi.org/10.1109/TNNLS.2020.2978386.
33. Chen, X., Liu, X., and Wu, J. (2020). GCN-BMP: investigating graph representation learning for DDI prediction task. Methods *179*, 47–54. https://doi.org/10.1016/j.ymeth.2020. 05.014.
34. Deac, A., Huang, Y.-H., Veličković, P., Liò, P., and Tang, J. (2019). Drug-drug adverse effect prediction with graph co-attention. Preprint at arXiv. https://doi.org/10.48550/arXiv.1905.00534.
35. Zitnik, M., and Zupan, B. (2016). Collective pairwise classification for multi-way analysis of disease and drug data. In Biocomputing 2016: Proceedings of the Pacific Symposium (World Scientific), pp. 81–92. https://doi.org/10.1142/9789814749411_0008.
36. Nickel, M., Tresp, V., and Kriegel, H.-P. (2011). A three-way model for collective learning on multi-relational data. Icml *11*, 3104482–3104584. https://doi.org/10.5555/3104482.3104584.
37. Lin, X., Quan, Z., Wang, Z.-J., Ma, T., and Zeng, X. (2020). KGNN: Knowledge Graph Neural Network for Drug-Drug Interaction Prediction. IJCAI *380*, 2739–2745. https://doi.org/10.24963/ijcai.2020/380.
38. Zitnik, M., Agrawal, M., and Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics *34*, i457–i466. https://doi.org/10.1093/bioinformatics/bty294.
39. Liu, S., Huang, Z., Qiu, Y., Chen, Y.-P.P., and Zhang, W. (2019). Structural network embedding using multi-modal deep auto-encoders for predicting drug-drug interactions. In 2019 IEEE International conference on bioinformatics and biomedicine (BIBM) (IEEE), pp. 445–450. https://doi.org/10.1109/BIBM47256.2019.8983337.
40. Wang, H., Lian, D., Zhang, Y., Qin, L., and Lin, X. (2020). Gognn: Graph of graphs neural network for predicting structured entity

interactions. Preprint at arXiv. https://doi.org/10.24963/ijcai.2020/183.

41. Wicha, S.G., Chen, C., Clewe, O., and Simonsson, U.S.H. (2017). A general pharmacodynamic interaction model identifies perpetrators and victims in drug interactions. Nat. Commun. *8*, 2129. https://doi.org/10.1038/s41467-017-01929-y.
42. Razek, A., Vietti, T., and Valeriote, F. (1974). Optimum time sequence for the administration of vincristine and cyclophosphamide *in vivo*. Cancer Res. *34*, 1857–1861.
43. Koizumi, W., Kurihara, M., Hasegawa, K., Chonan, A., Kubo, Y., Maekawa, R., Iwasaki, R., Sasai, T., Fukuyama, Y., Ishikawa, K., et al. (2004). Sequence-dependence of cisplatin and 5-fluorouracil in advanced and recurrent gastric cancer. Oncol. Rep. *12*, 557–561. https://doi.org/10.3892/or.12.3.557.
44. Feng, Y.-Y., Yu, H., Feng, Y.-H., and Shi, J.-Y. (2022). Directed graph attention networks for predicting asymmetric drug-drug interactions. Briefings Bioinf. *23*, bbac151. https://doi.org/10.1093/bib/bbac151.
45. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. (2017). Graph attention networks. Preprint at arXiv. https://doi.org/10.48550/arXiv.1710.10903.
46. Deng, Z., Xu, J., Feng, Y., Dong, L., and Zhang, Y. (2024). MAVGAE: a multimodal framework for predicting asymmetric drug-drug interactions based on variational graph autoencoder. Comput. Methods Biomech. Biomed. Eng. 1–13. https://doi.org/10.1080/10255842.2024.2311315.
47. Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. Preprint at arXiv. https://doi.org/10.48550/arXiv.1412.6980.
48. Feng, Y.-H., Zhang, S.-W., and Shi, J.-Y. (2020). DPDDI: a deep predictor for drug-drug interactions. BMC Bioinf. *21*, 419–515. https://doi.org/10.1186/s12859-020-03724-x.
49. Van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*. https://doi.org/10.48550/arXiv.2108.01301.
50. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. *46*, D1074–D1082. https://doi.org/10.1093/nar/gkx1037.
51. Weininger, D. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J. Chem. Inf. Comput. Sci. *28*, 31–36. https://doi.org/10.1021/ci00057a005.
52. Yang, J., Cai, Y., Zhao, K., Xie, H., and Chen, X. (2022). Concepts and applications of chemical fingerprint for hit and lead screening. Drug Discov. Today *27*, 103356. https://doi.org/10.1016/j.drudis.2022.103356.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Drugbank	Drugbank	RRID:SCR_002700
Asymmetric-DDI	This paper	https://github.com/Wzew5Lp/DRGATAN
Software and algorithms		
Python 3.8	Anaconda	https://www.anaconda.com
RDKit 2023.3.1	RDKit	https://www.rdkit.org/
Python-Geometric 2.2.0	Python-Geometric 2.2.0	https://pytorch-geometric.readthedocs.io/en/latest/
Pytorch 1.12.1	Pytorch	https://pytorch.org/
DRGATAN	This paper	https://github.com/Wzew5Lp/DRGATAN

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jin Li (lijin@ynu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- This paper analyzes existing, publicly available data. The data for drugs are available from https://www.drugbank.ca/. The data is publicly accessible in https://github.com/Wzew5Lp/DRGATAN/tree/master/my_dataset.
- The source code of the DRGATAN is available at https://github.com/Wzew5Lp/DRGATAN and is publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The dataset for predicting asymmetric drug-drug interactions can be downloaded from DrugBank (https://www.drugbank.ca/). DrugBank is a comprehensive bioinformatics and cheminformatics database developed by the University of Alberta, which combines detailed drug data with extensive drug target information. This dataset is an online database focused on drug and drug-related information, including a wide range of drug details such as chemical structures, pharmacological properties, pharmacokinetics, drug interactions, etc. We parsed approved small molecule drugs with asymmetric interaction information as the dataset for predicting asymmetric drug-drug interactions. After removing drugs with missing or erroneous SMILES strings, we obtained a final dataset comprising 1876 small molecule drugs and 218917 asymmetric interactions. For specific experimental details, please refer to the experimental description in the Results section.

METHOD DETAILS

Problem definition

The definition of a multi-relational directed graph of drug interactions based on asymmetric interactions between drugs is defined as $\mathcal{G} = \{(i,r,j)|i,j \in D, r \in R\}$, where D is the set of drugs and R is the set of asymmetric interaction types (directed edge sets). Each triple (i,r,j) represents an asymmetric drug interaction r between drug i and drug j. For nodes, $i,j \in D$, (i,j) represents a directed edge from i to j, that is, an asymmetric drug interaction r. Specifically, drug i affects drug j, which can be represented by $i \rightarrow j$. We treat asymmetric interaction as a directed relation. Under different directed relations, the role of drugs may change. Henceforth, we refer to i as the relation source drug and j as the relation target drug.

Given a known multi-relational directed graph $\mathcal{G}$ and the initial feature matrix $\mathbf{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{N \times F}$ of drugs, our main goal is to learn a prediction function $\widehat{y}_{i,j} = \mathcal{F}(i,j|\mathcal{P},\mathcal{G},\mathbf{X})$, predicting whether there is an asymmetric interaction between two drugs, where $\widehat{y}_{i,j}$ represents the probability that drug i and drug j have an asymmetric interaction of the type $i \rightarrow j$, and $\mathcal{P}$ represents the model parameters of the function $\mathcal{F}$. This task is to predict whether there is an asymmetric interaction between drugs. Specifically, it predicts whether the directed edge between drug i and drug j is $i \rightarrow j$, $j \rightarrow i$, or no directed edge.

Drug structure similarity features

The drug data information comes from DrugBank.[50] DrugBank is a comprehensive biological database with detailed information on drugs. Clinical data on drugs include drug-drug interactions, drug-food interactions, side effects, etc.

SMILES(Simplified Molecular Input Line Entry Specification) string[51] is a specification that uses ASCII strings to explicitly describe the molecular structure. SMILES strings can be interpreted by molecular editing software and used to generate 2D molecular structures. Figure S1 shows the molecular structure of Morphine along with its corresponding SMILES string.

First, molecular fingerprints[52](such as Morgan fingerprints), corresponding to 1876 small molecule drugs in our dataset are obtained based on their SMILES string. Subsequently, the structural similarity between each drug and the remaining 1875 drugs is calculated using the Tanimoto coefficient to derive the drug's SSP feature vector. Using PCA to reduce dimensionality, the low-dimensional vector of the SSP of each drug is obtained as the initial feature of the drug. The specific process is shown in Figure S2.

Relation role encoding

In an asymmetric interaction relation, the two drugs are not equivalent, as one drug affects the pharmacological effects of the other. We learn two role-specific representations under asymmetric relations, namely the relation source role and the relation target role.

In the multi-relational directed graph $\mathcal{G} = \{(i,r,j)|i,j \in D, r \in R\}$, two types of neighborhoods are defined for the relation source role and relation target role of node i. One is the first-order (out-degree) outgoing neighbor of its relation source role i as an attacker. The other is the first-order (in-degree) incoming neighbor of its relation target role i as a victim. Then, RGAT is employed to aggregate neighboring nodes and relation information using relation graph attention mechanisms, facilitating the learning of relation source and target role embeddings. The overall structure of RGAT is shown in Figure S3.

Use RGAT to aggregate the neighbor node information of node i through the relation graph attention mechanism. Taking i as an example of the relation source role, the attention coefficient $\alpha_{i,j}$ with its outgoing neighboring nodes under different relation r is defined as shown in Equation 1.

$$\alpha_{i,j} = \frac{\exp\left(LeakyReLU\left(\boldsymbol{x}_i^T \mathbf{W}_r \boldsymbol{q}^r + \boldsymbol{x}_j^T \mathbf{W}_r \boldsymbol{k}^r\right)\right)}{\sum_{n \in \mathcal{N}_o^r(i)} \exp\left(LeakyReLU\left(\boldsymbol{x}_i^T \mathbf{W}_r \boldsymbol{q}^r + \boldsymbol{x}_j^T \mathbf{W}_r \boldsymbol{k}^r\right)\right)} \quad \text{(Equation 1)}$$

Where $\mathcal{N}_o^r(i)$ is the first-order outgoing neighbor node set of node i under the relation $r \in R$. $\boldsymbol{x}_i \in \mathbb{R}^F$ is the initial feature of node i. $\mathbf{W}_r \in \mathbb{R}^{F \times \acute{F}}$ is the learnable parameter matrix for each relation r. $\boldsymbol{q}^r \in \mathbb{R}^{\acute{F}}$ and $\boldsymbol{k}^r \in \mathbb{R}^{\acute{F}}$ are query and key cores, respectively, which map the intermediate representation. LeakyReLU is a non-linear activation function. The attention coefficient $\alpha_{i,j}$ aggregates different neighbor nodes j and different relation importance information of node i under the same role information.

Finally, the relation source role embedding and the relation target role embedding are obtained through a combination of the neighborhood aggregation step and the attention mechanism, defined by Equations 2 and 3.

$$\boldsymbol{o}_i = \sigma\left(\sum_{r \in R} \sum_{j \in \mathcal{N}_o^r(i)} \alpha_{i,j} \mathbf{W}_r^T \boldsymbol{x}_j\right) \quad \text{(Equation 2)}$$

$$\boldsymbol{m}_i = \sigma\left(\sum_{r \in R} \sum_{j \in \mathcal{N}_m^r(i)} \alpha_{i,j} \mathbf{W}_r^T \boldsymbol{x}_j\right) \quad \text{(Equation 3)}$$

Where $\mathcal{N}_m^r(i)$ is the first-order incoming neighbor node set of node i under the relation $r \in R$. $\boldsymbol{o}_i \in \mathbb{R}^{\acute{F}}$ is the relation source role embedding of node i. $\boldsymbol{m}_i \in \mathbb{R}^{\acute{F}}$ is the relation target role embedding of node i. The process of embedding the source role and target role of the relation is shown in Figure S4.

Relation self-role embedding encodes both chemical structure information and network structure information. For the asymmetric action $i \rightarrow j$, the relation self-role $\boldsymbol{c}_i \in \mathbb{R}^F$ is defined as shown in Equation 4.

$$\boldsymbol{c}_i = \frac{1}{|\mathcal{N}(i)|} \sum_{r \in R} \sum_{n \in \mathcal{N}_r(i)} \left(\alpha_n \mathbf{M}_r \boldsymbol{x}_n + \boldsymbol{x}_n\right) \quad \text{(Equation 4)}$$

Where $\mathcal{N}(i)$ is the set of all first-order neighbors of node i and $j \notin \mathcal{N}(i)$. $\mathbf{M}_r \in \mathbb{R}^{F \times F}$ is the parameter matrix of the corresponding relation. α_n is the updatable weight parameter for each node.

Invasiveness and vulnerability

In social network analysis, "centrality" is a measure commonly used to assess the importance or influence of nodes within a network. The most direct measure of centrality is Degree Centrality, which indicates that a node is more important if it has a higher degree in the network.

Building on this concept, we hypothesize that in a directed DDI network, drugs with a higher number of interactions are more likely to interact with a greater number of drugs. In short, the greater the in-degree of node i, the greater its vulnerability, and the greater the out-degree of node i, the greater its invasiveness. We refer to these two situations as the invasiveness of the relation source role and the vulnerability of the relation target role.

We use degree centrality to reflect the correlation between nodes. Considering that the aggregation process of RGAT also takes into account the degree of nodes. Hence, it is combined with degree centrality to provide additional information. Define two scalars a_i and v_i to represent invasiveness and vulnerability, respectively. a_i represents the invasiveness of the relation source role embedding of node i. v_i represents the vulnerability of the relation target role embedding of node i. Their calculation methods are shown in Equations 5 and 6.

$$a_i = \frac{|\mathcal{N}_o(i)|}{|D| - 1} + p \quad \text{(Equation 5)}$$

$$v_i = \frac{|\mathcal{N}_m(i)|}{|D| - 1} + q \quad \text{(Equation 6)}$$

Where $\mathcal{N}_o(i)$ is the number of first-order out-degree neighboring nodes of relation source role i. $\mathcal{N}_m(i)$ is the number of first-order in-degree neighboring nodes of relation target role i. $|D|$ is the number of all nodes in the directed DDI network. $p \in \mathbb{R}$ is the last element value of $\boldsymbol{o}_i$. $q \in \mathbb{R}$ is the last element value of $\boldsymbol{m}_i$.

Prediction and loss

When the influence of i on j exists, we represent the asymmetric interaction by considering i as the relation source role and j as the relation target role as the central nodes. We obtain the possibility of asymmetric interaction from different central nodes. Finally, we obtain the final possibility $\hat{y}_{i,j}$ by combining these two types of possibility information through summation. Its definition is shown in Equation 7.

$$\hat{y}_{i,j} = \sigma\left(\lambda_1 * \acute{\boldsymbol{c}}_j \boldsymbol{o}_i^T + \lambda_2 * a_i + \lambda_1 * \acute{\boldsymbol{c}}_i \boldsymbol{m}_j^T + \lambda_2 * v_j\right) \quad \text{(Equation 7)}$$

Where $\acute{\boldsymbol{c}}_j = \acute{\mathbf{W}}_o^T \boldsymbol{c}_j$ and $\acute{c}_i = \acute{\mathbf{W}}_m^T c_i$, $\acute{\mathbf{W}}_o^T \in \mathbb{R}^{\acute{F} \times F}$ and $\acute{\mathbf{W}}_m^T \in \mathbb{R}^{\acute{F} \times F}$ are alignment matrices. $\boldsymbol{o}_i$ is the relation source role embedding of node i. $\boldsymbol{m}_j$ is the relation target role embedding of node j. σ is the Sigmoid function. $\lambda_1 + \lambda_2 = 0.5$ is used to constrain the probability value. λ_1 controls the similarity of relation role embedding between constrained nodes. λ_2 controls the invasiveness and vulnerability of constrained nodes under different relation role embeddings.

Our task is to determine whether there is an asymmetric interaction between drugs i and j. Given that, there exists an asymmetric interaction from $i \rightarrow j$ between the drug pair (i, j) when $\hat{y}_{i,j} \geq 0.5$.

Finally, the binary cross-entropy loss function for the overall model framework training is shown in Equation 8:

$$\mathcal{L} = \frac{1}{N} \sum_{(i,r,j) \in Y} - y_{i,j} \log\left(\hat{y}_{i,j}\right) - \left(1 - y_{i,j}\right) \log\left(1 - \hat{y}_{i,j}\right) \quad \text{(Equation 8)}$$

Where $y_{i,j}$ is the ground truth label. $y_{i,j} = 1$ means that there is an asymmetric interaction between the nodes, and $y_{i,j} = 0$ means that there is no asymmetric interaction between the nodes. Y is the set of drug pairs and their interactions.

QUANTIFICATION AND STATISTICAL ANALYSIS

We extracted six different drug molecular fingerprint features to study the performance of the DRGATAN model. These drug molecular fingerprints include MACCS, Atom Pair, ECFP4, ECFP6, RDKit, and Topological Torsion Fingerprint. Among these drug molecular fingerprint features, based on Structural Similarity Profile (SSP), achieved the best performance under the same feature conditions. Experimental results indicate that using SSP to generate drug features has better performance in asymmetric DDI prediction tasks, with better representational capability to improve predictive performance. We also compared the experimental results with other advanced methods, and DRGATAN achieved satisfactory results. To demonstrate the effectiveness of relationship information and attention mechanisms in modeling asymmetric information, we visualized drug feature information in different scenarios for validation. Additionally, we predicted 20 potential asymmetric drug interactions to analyze the practical effectiveness of the model. The statistical results and their explanations are detailed in the main text of this paper.