Zhe JIN, Yin Zhang, Jiaxu MIAO, Yi YANG, Yueting ZHUANG, Yunhe PAN, 2023. A knowledge-guided and traditional Chinese medicine informed approach for herb recommendation. *Frontiers of Information Technology & Electronic Engineering*, 24(10):1416-1429. https://doi.org/10.1631/FITEE.2200662

A knowledge-guided and traditional Chinese medicine informed approach for herb recommendation

Key words: Traditional Chinese medicine; Herb recommendation; Knowledge graph; Graph attention network

Corresponding author: Yin ZHANG
E-mail: yinzh@zju.edu.cn
ORCID: https://orcid.org/0000-0001-6986-4227

Background

Traditionally, traditional Chinese medicine (TCM) consultation has been mainly offline. Doctors obtain the symptoms of patients by four basic diagnostic methods called *wang-wen-wen-qie* and develop a prescription for treatment. The whole process is based mainly on the doctor's experience in treating patients and clinical practice, which is highly subjective. Several studies have tried to learn such prescribing patterns and knowledge through some machine learning algorithms with incorporating some external TCM knowledge. However, the introduction of this knowledge remains inadequate. It remains a challenge to introduce sufficient external knowledge in a suitable model.

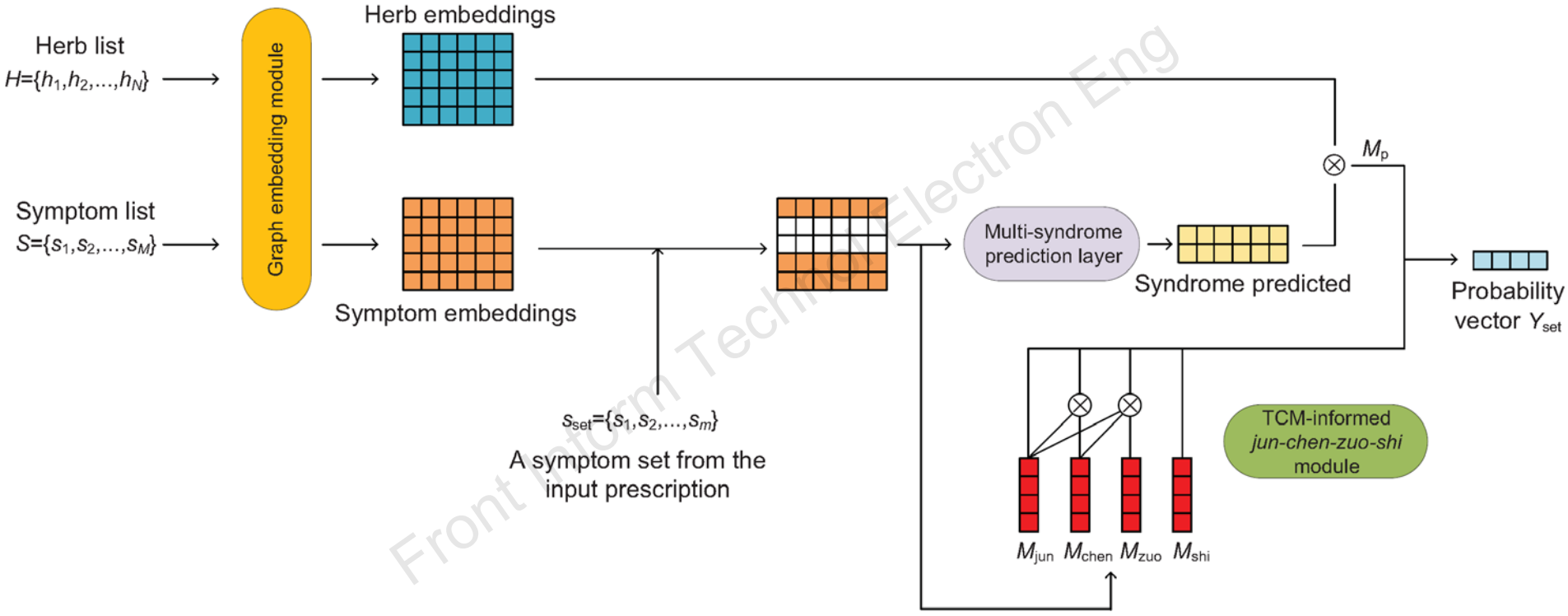
Main idea

1. We construct a large-scale TCM knowledge graph, TCMKG, which consists of herbs, prescriptions, pieces, diseases, symptoms, syndrome, tropism, properties, flavor, and functions, and covers most of the common knowledge in TCM.

2. We propose a knowledge-guided and TCM-informed approach combining path interactions and co-occurrence relationships from the knowledge graph, multi-syndrome prediction, and the TCM-informed principle of prescription composition.

3. Experimental results on a TCM prescription dataset show that our proposed approach has better performance than state-of-the-art TCM herb recommendation models.

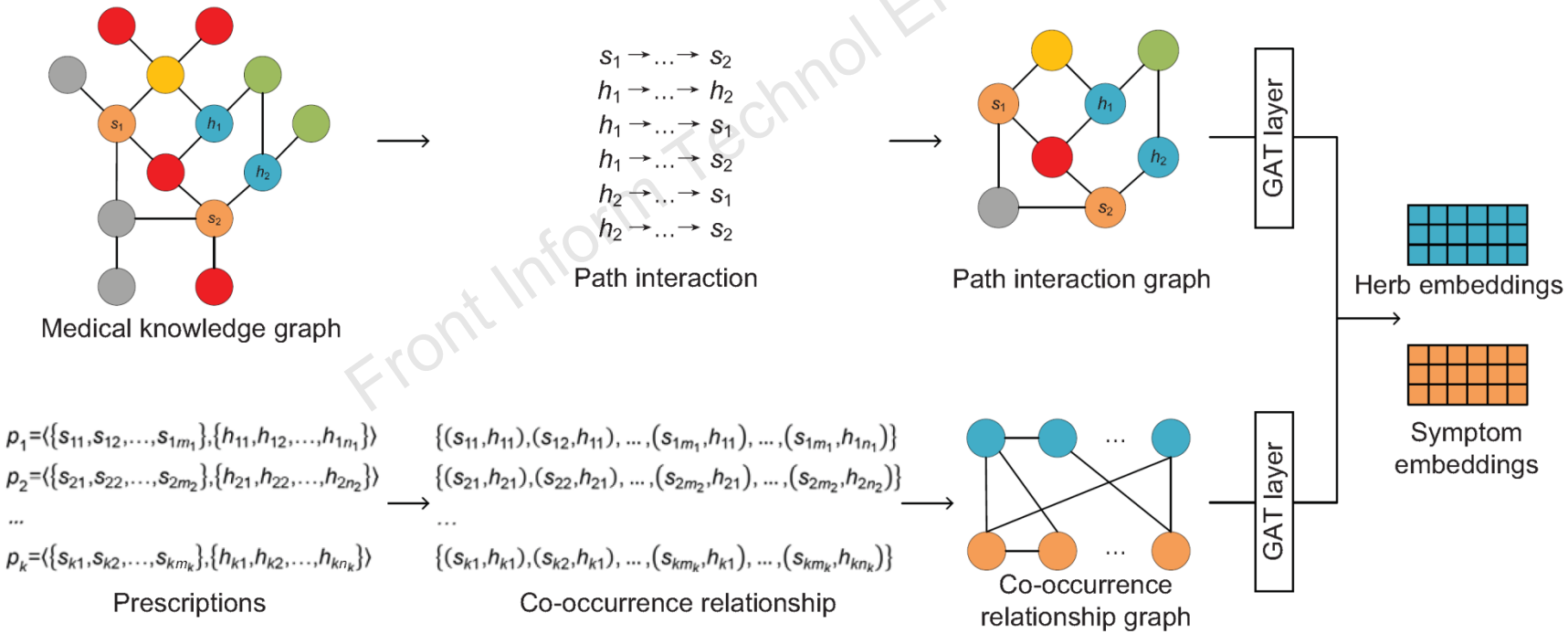
Framework



The overall architecture of our proposed approach, MKMI-GAT

Model

1. We construct a TCM knowledge graph called TCMKG by Neo4j. Based on TCMKG, we propose a graph embedding module by constructing two graphs and using a graph attention network (GAT) to learn discriminative feature embeddings of herbs and symptoms.



The graph embedding module

Model

2. As one syndrome may encompass multiple symptoms and one symptom may appear in multiple syndromes, we propose a multi-syndrome prediction layer to learn the representation of potential syndromes.

3. We add a TCM-informed *jun-chen-zuo-shi* module to add constraints of TCM-informed principles in deep learning.

$$\hat{M}_{\text{jun}} = M'_{\text{jun}} \qquad \hat{M}_{\text{chen}} = W_{\text{chen}} M'_{\text{jun}} + M'_{\text{chen}}$$

$$\hat{M}_{\text{zuo}} = W_{\text{zuo}}(M'_{\text{jun}} + M'_{\text{chen}}) + M'_{\text{zuo}} \qquad \hat{M}_{\text{shi}} = M'_{\text{shi}}$$

Major results

Table 1 Performance evaluation of different models

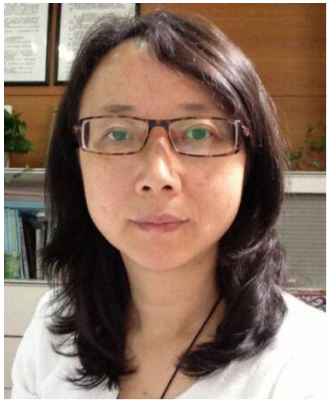
Model	Precision			Recall			F1-Score		
	P@5	*P*@10	*P*@20	*R*@5	*R*@10	*R*@20	F1@5	F1@10	F1@20
GATConv (Veličković et al., 2018)	0.192	0.151	0.114	0.132	0.183	0.288	0.156	0.165	0.163
GCNConv (Kipf and Welling, 2017)	0.211	0.169	0.124	0.138	0.218	0.320	0.167	0.190	0.178
HC-KGETM (Wang XY et al., 2019)	0.181	0.142	0.106	0.120	0.199	0.291	0.145	0.166	0.155
SMGCN (Jin et al., 2020)	0.193	0.157	0.115	0.127	0.203	0.303	0.153	0.177	0.167
KDHR (Yang et al., 2022)	0.237	0.188	0.140	0.182	0.283	0.416	0.206	0.226	0.210
MKMI-GAT	**0.288**	**0.229**	**0.169**	**0.210**	**0.327**	**0.474**	**0.243**	**0.269**	**0.249**

The best values are in bold

Conclusions

We propose a knowledge-guided and TCM-informed approach based on GATs for herb recommendation. We construct a large-scale TCM knowledge graph, TCMKG, which covers most of the common knowledge in the TCM field. The TCMKG path interaction graph and co-occurrence relationship graph based on the treatment relationship from prescriptions are trained by GATs and output the discriminative representation of herbs and symptoms. After the fusion of these two representations, the multi-syndrome prediction layer will generate a possible representation of syndromes through multi-head attention. The model also introduces the TCM-informed *jun-chen-zuo-shi* module, which provides knowledge about the principles of prescription composition.

Zhe JIN received his PhD degree from College of Computer Science and Technology, Zhejiang University, Hangzhou, China in 2023. He has been involved in the design and construction of the Professional Knowledge Service System for Chinese Herbal Medicine from the China Knowledge Centre for Engineering Sciences and Technology (CKCEST). His current research interests include knowledge graph and deep learning.



Yin ZHANG received her PhD degree in computer science from Zhejiang University in 1999. Currently, she is an associate professor with College of Computer Science and Technology, Zhejiang University. Her research interests include mainly data mining and knowledge discovery, medical informatics, multimedia information processing, pattern recognition, and knowledge engineering.