

• 综 述 •

基于重叠社区的“方-药”网络经方配伍规律分析

梁力伟¹, 丁长松^{1*}, 黄辛迪¹, 梁 昊², 蔡 雄³

1. 湖南中医药大学信息科学与工程学院, 湖南 长沙 410208

2. 湖南中医药大学 中医诊断研究所, 湖南 长沙 410208

3. 湖南中医药大学 科技创新中心, 湖南 长沙 410208

摘 要: 针对传统数据挖掘技术难以揭示海量方剂配伍复杂性的问题, 借鉴复杂网络思想构建“方-药”网络并基于重叠社区分析其组方配伍规律。首先将方剂内包含的药物抽象为节点、药间关系抽象为边、药物共现频次表示为边的权重, 构建“方-药”网络模型, 引入“药物影响力”的概念识别网络中的重要药物节点; 其次提出一种基于重叠社区的“方-药”网络划分方法进行社区发现; 最后分析划分后社区中蕴含的组方配伍规律。在基于《伤寒论》112 首方剂构建的模型中, 药物影响力大的药物符合张仲景“寒温并用、辛甘苦同行”的用药原则, 结合 FCM (fuzzy C-means) 算法可发现隐藏在《伤寒论》整体模型中的常用药组; 在基于《伤寒论》少阴病证构建的模型中, 药物影响力大的药物符合少阴病证“温经扶阳、利水渗湿”的用药特点, 并得到了分类明显的药物配伍社团。“方-药”网络模型中蕴含着高质量的重叠社区结构, 分析社区中蕴含的内在规律可为方剂组方配伍及用药规律的探究提供新的依据。

关键词: 复杂网络; “方-药”网络; 药物影响力; 重叠社区; 用药规律

中图分类号: R285 **文献标志码:** A **文章编号:** 0253-2670(2020)02-0496-11

DOI: 10.7501/j.issn.0253-2670.2020.02.029

Combination rule analysis of classic formulae in “formulae-herb” network based on overlapping community

LIANG Li-wei¹, DING Chang-song¹, HUANG Xin-di¹, LIANG Hao², CAI Xiong³

1. School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China

2. Institute of TCM Diagnostics, Hunan University of Chinese Medicine, Changsha 410208, China

3. Institute of Innovation and Applied Research in Chinese Medicine, Hunan University of Chinese Medicine, Changsha 410208, China

Abstract: The compatibility of Chinese materia medica (CMM) studies the relations between herbs and their interacting effects. With complex medications for individuals and massive CMM data, it is difficult for traditional data mining methods to reveal the inherent rules in compatibility analysis. This paper with the theory of complex network constructs a “formulae-herb” network to analyze herb combinations based on the overlapping communities for internal rules in CMM classic formulae. Based on the complex network, the formulae can be abstracted to a graph and represented by nodes and edges, herbs as nodes, and the relationships of herbs as edges. The edges were weighted according to the frequency of co-occurrence of two herbs. “Formulae-herb” network was then constructed based on all formulae in the study. A definition “herb influence” was introduced into the network to identify important herb nodes, with which a method of “formulae-herb” network division was introduced based on overlapping communities to discover the structure of the herb relations. Finally, the inherent laws in the community were discussed. A “formulae-herb” network was constructed based on 112 formulae from *Treatise on Cold Damage*. Herbs with high influence are in line with Zhang Zhong-Jing's

收稿日期: 2019-10-21

基金项目: 国家重点研发项目 (2017YFC1703306); 湖南省自然科学基金资助项目 (2018JJ2301); 湖南省重点科技计划项目 (2017SK2111); 湖南省教育厅科学研究项目 (18A227); 湖南省教育厅科学研究项目 (18C0380); 湖南省教育厅科学研究项目 (18K070); 湖南省中医药科研计划重点项目 (2020002)

作者简介: 梁力伟 (1995—), 研究生, 研究方向为大数据与计算中医学。Tel/Fax: 13971238331 E-mail: 215799292@qq.com

***通信作者** 丁长松 Tel/Fax: (0731)88458173 E-mail: 15208250@qq.com

medication principle of “combining cold and warm, use of pungent, sweet and bitter together”. This paper also incorporates FCM (fuzzy C-means) algorithm to discover the commonly used herb groups hidden in the medication. A model of treating Shaoyin syndrom was constructed as an application. Herbs with high influence were in line with the treating method of “warming meridian and reinforcing yang, promoting urination and draining dampness”. The result showed that multiple herb groups were clearly classified. “Formulae-herb” network model showed high quality overlapping community structures, the analysis of inherent laws in the community can provide basis for the exploration of compatibility rules of herbs and internal rules in formulae regularities.

Key words: complex network; “formulae-herb” network; drug impact; overlapping community; medication rules

方剂是在中医药理论的指导下,依据组方原则选择合适的药物配伍而成^[1],具有以“方-药-证”为核心的多维结构^[2],深入探究方剂组方配伍规律对于阐明和发展中医药理论意义重大。传统的中医药数据挖掘往往通过对方药数据进行特征提取,结合分类^[3]、聚类^[4]、关联规则^[5]等技术分析,这种通过统计归纳处理对象的方法偏向于发掘数据表层规律,缺乏对系统内在机制或内涵的深入讨论^[6],难以揭示隐含在方剂组方配伍中的深层规律。

复杂网络以网络化建模形式研究对象间存在的复杂现象^[7],并常结合社区发现分析其中的社区结构、社区特征及社区间关系^[8]。方剂本身是由多味药构成且一味药可能存在于多个方剂中,这样便产生了方剂间药物相交的情况,若将其反映于真实“方-药”网络,不同的方剂可视为不同的社区,方剂间药物的相交便可视为社区的重叠,这为“方-药”网络中重叠社区的发现提供了基础。故本研究借助复杂网络的思想以单味药为节点、药间关系为边、药物共现频次为边的权重,构建“方-药”网络模型,引入“药物影响力”的概念识别网络中的重要药物节点,并提出一种适用于“方-药”网络模型的重叠社区划分方法对“方-药”网络中存在的深层关系进行分析,以期为中医临床辨证论治、遣药组方提供依据。

1 方法

1.1 数据来源与整理

本文使用的数据均来自于全国中医药行业高等教育“十三五”规划教材《伤寒论选读》新世纪第4版(王庆国主编,2016年版,以下简称为“教材”)。收集统计了教材中涉及的112首方剂信息,包括构建“方-药”网络模型需要的方剂名称与药物名称(未载入缺药物组成的禹余粮丸方)。依据全国中医药行业高等教育“十三五”规划教材《中药学》(钟赣生主编,2016年版),对112首方剂中的药物名称进行了规范统一,如淡豆豉、香豉统一为香豉、代赭统一为代赭石、麦冬统一为麦门冬等,规范整理后共涉及92味药物,部分方剂信息见表1。

表1 规范后《伤寒论选读》方剂示例

Table 1 Examples of prescription for *Selected Reading of the Treatise on Cold Damage* after specification

编号	方剂名称	药物组成
1	桂枝汤	桂枝、芍药、炙甘草、生姜、大枣
2	麻黄汤	麻黄、桂枝、炙甘草、杏仁
3	葛根汤	葛根、麻黄、桂枝、生姜、炙甘草、芍药、大枣
4	大青龙汤	麻黄、桂枝、炙甘草、杏仁、生姜、大枣、生石膏

1.2 “方-药”网络模型构建

“方-药”网络模型依据教材中各方剂药物组成建立,描述方剂中药物间存在的相互关系。以方剂中包含的所有药物为节点,若2个药物出现在同一个方剂中,则在2个药物之间建立1条边,其权重为2个药物共现的频次。在模型构建过程中,本文引入“药物影响力”这一指标识别模型中的重要药物节点。该指标从药物节点间连边出发,通过药物节点的权值比之和来反映药物节点重要性,具体公式如下。

$$I_{\text{score}(v)} = \sum_{v=1}^{d_v} \left(\frac{W_{vu}}{\sum_{u=1}^{d_u} W_{uu'}} \right)$$

$I_{\text{score}(v)}$ 表示药物节点 v 的药物影响力, d_v 与 d_u 分别表示药物节点 v 和药物节点 u 的度; W_{vu} 表示药物节点 v 和其邻节点 u 之间的权值, $W_{uu'}$ 表示药物节点 u 和其邻节点 u' 之间的权值; W_{vu} 与 $\sum_{u=1}^{d_u} W_{uu'}$ 的比值表示药物节点 v 的权值比

1.3 基于重叠社区的“方-药”网络划分

1.3.1 相关概念

(1) 药物对社区贡献度:表示节点与社区之间的关联强度,具体公式如下。

$$C(u, C) = \frac{\sum_{v \in C} W_{uv}}{\sum_{u=1}^{d_u} W_{uu'}}$$

$C(u, C)$ 表示节点 u 与社区 C 间的“药物对社区贡献度”，

$\sum_{v \in C} W_{uv}$ 表示节点 u 与社区 C 内节点的总权值和， $\sum_{u=1}^{d_u} W_{uu'}$ 表示

节点 u 与“方-药”网络中所有节点的权值和

(2) 自适应函数：表示社区内节点的紧密程度，具体公式如下。

$$f_g(C) = \frac{k_{in}^g}{(k_{in}^g + k_{out}^g)^\alpha}$$

借鉴 Lancichinetti 等^[9]提出的自适应函数并结合“方-药”网络的实际结构，在本文中 k_{in}^g 表示社区 C 内部节点间的权值和，而 k_{out}^g 表示社区 C 的内部节点与剩余外部节点间的权值和。参数 α 是一个

为正的实数，范围 $[0, 1]$ ，用来控制社区发现规模。

1.3.2 “方-药”网络重叠社区划分 方剂间药物的相交导致真实“方-药”网络中存在着大量的重叠药物节点，借鉴於志勇等^[10]提出的 i-SE OCD 算法 (influence seeds extension overlapping community detection) 并结合构建的“方-药”网络模型，本文提出了一种适用于“方-药”网络模型的重叠社区划分方法。该方法的核心思想是在基于种子扩展的局部社区发现算法^[9-12]基础上，优化药物种子节点的选取实现“方-药”网络的重叠社区划分，共包括 3 部分，其中第 1 部分为核心药物社区的发现，第 2 部分为社区扩展，第 3 部分为社区优化，总流程见图 1。

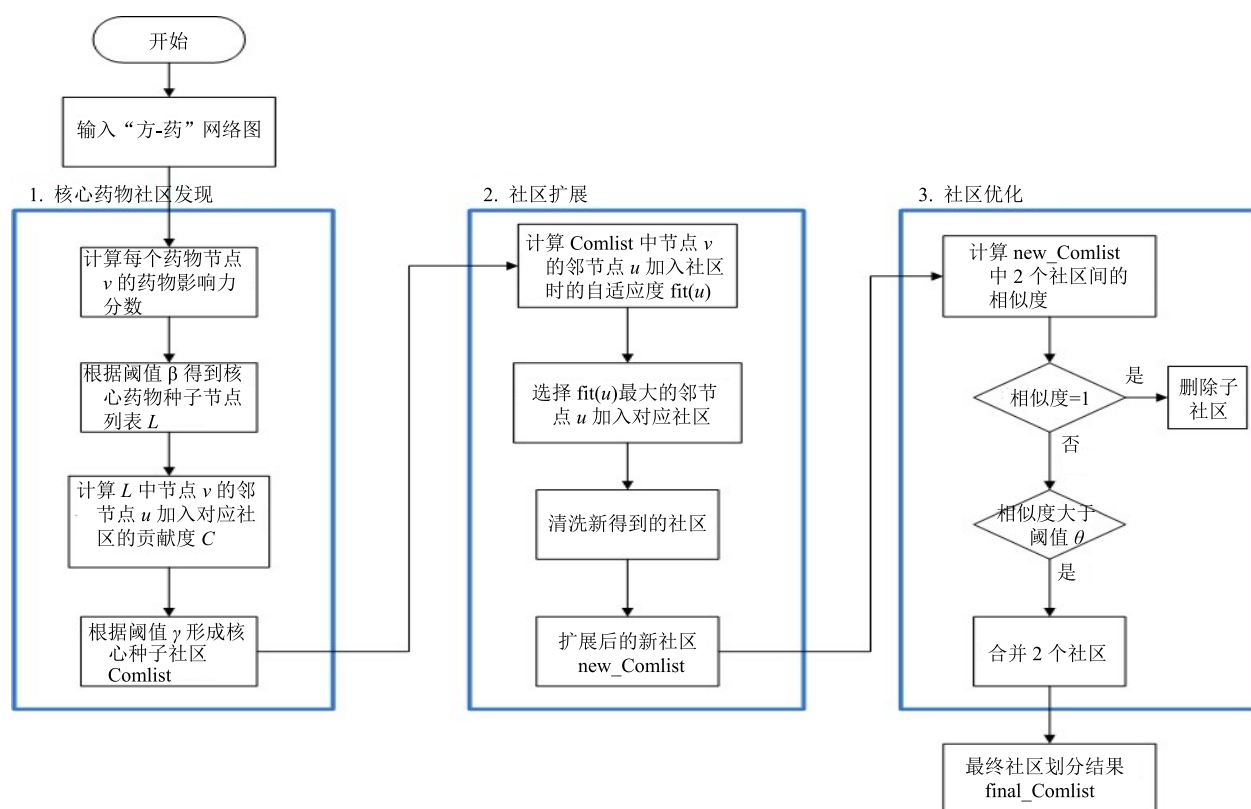


图 1 “方-药”网络重叠社区划分流程

Fig. 1 Division process of “formulae-herb” network overlapping community

(1) 核心药物社区发现：核心药物社区发现过程的总体思想是引入“药物影响力”及阈值 β 实现核心药物种子节点的选取；引入“药物对社区贡献度”及阈值 γ 实现核心药物社区的发现。具体步骤如下：①计算“方-药”网络中所有药物节点 v 的药物影响力 $I_{score}(v)$ 。②选择 1 个药物节点 v ，统计 v 的邻节点总个数 (N_{num})。③比较药物节点 v 的药物影

影响力 $I_{score}(v)$ 与其邻节点 u 的药物影响力 $I_{score}(u)$ ，统计小于 $I_{score}(v)$ 的邻节点个数 (L_{num})。④计算 L_{num} 与 N_{num} 的比值，若比值超过阈值 β ，取 v 为核心药物种子节点并作为一个独立的社区。⑤对于剩余药物节点重复步骤②~④，形成核心药物种子列表 L 。⑥任选 L 中的一个核心药物节点 v' ，取 v' 的一个邻节点 u' 。⑦计算 u' 加入对应社区时的药物对社区贡

献度 C ，若 C 大于阈值 γ 则将 u' 加入到社区中。⑧对 v' 的剩余邻节点重复步骤⑦，得到一个稳定紧密的核心药物社区。⑨对 L 中的剩余核心药物节点重复步骤⑥~⑧，最终得到多个稳定紧密的核心药物社区 C 。

(2) 社区扩展：得到稳定紧密的核心药物社区 C 后，本文通过迭代添加其邻节点到社区实现社区的扩展。用于社区扩展的方法包括适应度函数方法^[9] (LFM 算法)、R 方法^[13] 和标签传播^[14] 等。由于 LFM 算法^[9] 时间复杂度较低且应用范围不受限，本文借鉴 LFM 算法并结合重定义的自适应函数实现了社区的扩展，具体步骤为：①任选核心药物社区 C 中的一个社区 C_1 ，取 C_1 中的一个药物节点 v_1 。②计算 v_1 的所有邻节点 u_1 分别加入 C_1 的节点适应度^[10]。③选择节点适应度最大且为正值的 u_1 加入 C_1 ，得到一个新的社区 C_2 。④重新计算 C_2 中药物节点的节点适应度，删除第 1 个适应度为负值的药物节点，得到一个新的社区 C_3 。⑤对 C_3 重复步骤④，直至清洗完所有适应度为负值的药物节点。⑥对 C_1 中的剩余药物节点重复步骤②~⑤，得到一个扩展的新社区。⑦对 C 中的剩余社区重复步骤①~⑥，最终得到扩展后的新社区 new_C 。

(3) 社区优化：为了从整体出发综合分析社区划分后的特性，本文通过引入“社区相似度”^[10] 及阈值 θ 实现社区分类结果的优化。具体步骤为：①选择扩展后的新社区 new_C 中任意 2 个社区 C_4 和 C_5 。②分别计算 C_4 、 C_5 的社区规模 L_1 、 L_2 ，以及 C_4 和 C_5 的社区间相似度 (SC)。③判断 SC 是否等于 1，若等于 1 则比较 L_1 和 L_2 ，若 $L_1 \leq L_2$ ，删除社区 C_4 ，否则删除社区 C_5 ，得到一个新的社区 new_C' 。④若 SC 不等于 1，判断 SC 是否大于阈值 θ ，若大于则合并 C_4 和 C_5 得到一个新的社区 C_6 ，并将 C_6 添加至 new_C 中形成一个新的社区 new_C' 。⑤对 new_C' 中的社区重复步骤①~④，最终得到社区划分结果 $final_C$ 。

2 结果

2.1 《伤寒论》整体模型构建及社团划分

2.1.1 模型构建 基于《伤寒论》中 112 首方剂构建的“方-药”网络模型 (简称“《伤寒论》整体模型”) 见图 2。其中节点的大小反映药物的药物影响力强弱、颜色代表药物的分类。图 2 共包含 92 个节点，代表 112 首方剂中的 92 味中药，587 条边代表 112 首方剂中存在的 587 种“中药-中药”间关系。

节点的大小反映出在《伤寒论》整体模型中药物影响力较大的 5 味药为炙甘草、桂枝、大枣、生姜、半夏，其中包含 2 味和中补虚药 (炙甘草和大枣)、2 味辛温解表药 (桂枝和生姜) 以及 1 味化痰药 (半夏)。依据药物影响力的强弱可分析出张仲景治疗伤寒疾病时多选解表祛邪药及化痰药，并佐以和中补虚药以护脾胃，体现了其治疗伤寒疾病用药“寒温并用、辛甘苦同行”的用药规律^[15]。除了生甘草、文蛤、妇人中葶 3 个单独中药节点外，剩余药物构成了一个大的连通图，反映出 112 首方剂中存在着较多的重叠药物节点。

2.1.2 社团划分 经过后续参数实验 (见“2.3.1”)，本文最终选取 $\alpha=0.6$ 、 $\beta=0.4$ 、 $\gamma=0.08$ 、 $\theta=0.97$ 进行了《伤寒论》整体模型社区划分，划分结果共包含 8 个社区，将其用 Cytoscape 软件进行可视化显示得到的最终结果见图 3。图 3 共包含 75 味药物，不同的节点颜色代表药物不同的社团分类，如粉色药物均属社团 2、绿色药物均属社团 3，白色药物则代表划分后社区中的重叠药物节点。分析图中药物的分布情况可以看出，虽然整个社区划分结果包含一个以桔梗、甘草、贝母、巴豆 4 味药组成的独立社区，但剩余 7 个社区中重叠节点数量过多，社团结构不清晰，为分析同社团内药物节点间的联系带来了难度。

为了进一步分析重叠节点中蕴含的配伍规律，对麻黄、桂枝、细辛等 62 个重叠药物节点，采用经典的 FCM 算法 (fuzzy C-means)^[16] 进行聚类划分。以 62 味药物分别对 8 个社区贡献的权值为输入，并通过后续的参数实验 (见“2.3.2”) 选取模糊加权指数 (m) = 2，聚类个数 (k) = 19，划分结果如表 2 所示。表 2 共包含 62 味药物，19 种药物分类结果。其中包括 10 种由 2 味或 2 味以上药物组成的药物分类 (类别 1~10)，以及分别由大枣、炙甘草、生姜等 9 味药物组成的 9 种药物分类 (单独分类节点)。分析在表 2 中药物的分类情况可以发现，FCM 算法倾向于将常用药组划分到同一类别，如类别 1 中的半夏辛散、性温属阳，黄芩苦降、性寒属阴，2 药配伍可用于治疗慢性胃炎、急性胃肠炎见心下痞满、干呕或呕吐、肠鸣下利等症^[17]；类别 3 中的大黄入里以除腑实，柴胡走表以清经凑之邪热，2 药同用具有较好的通腑泄热、疏肝利胆的作用^[18]；黄连苦寒治热偏于清，且清热之中能止血，与大黄合用既能增强清泻作用，又能增强止血不留瘀之效^[19]；类别 7 中生石膏辛寒、细辛辛温，2 药配伍有“火郁散之”

之意,可用于熄风通络^[20],此外,临床上也常将当归与细辛配伍使用来温养肝血、通经和脉、润肝燥、止肝痛^[21];类别 10 中的枳实与厚朴均为苦味,其中枳实善破气兼泻火,厚朴下气除满兼燥湿,二者配伍使用可治气机郁结于胃肠之证^[22]。

除此之外,结合药物在《伤寒论》112 首方中的出现次数(部分统计信息见表 3)分析可得,通过 FCM 算法归类得到的单独分类节点,如炙甘草、桂枝、大枣等药物,在实际《伤寒论》112 首方中的出现次数均比较高。如炙甘草在 112 首方中的出现次数高达 64 次,桂枝与大枣 38 次,生姜 36 次等。这一情况符合张仲景的实际用药原则,仲景治疗伤寒疾病时将调和营卫之治法贯穿始终,同时顾护中焦脾土使五脏调和^[15],所以大枣、炙甘草、生姜、桂枝、芍药 5 味药出现的频数很高。除此之外,结合剩余的麻黄、茯苓、人参、干姜 4 味药,以上 9 味药又分别可以组成四君子汤、四逆汤、苓桂术甘汤等方剂,以实现保胃气顾护津液、调节水气出入的功效^[23]。故在实际重叠药物节点分类过程中,与其他药物节点不同,炙甘草、桂枝、大枣等药物具有高频数及高连接性,易被划分为独立节点。

综上所述,运用本文提出的“方-药”网络重叠社区划分方法并结合 FCM 算法,可发现隐藏在《伤寒论》整体模型中的常用药组。但由于《伤寒论》总 112 首方剂包含的桂枝汤及其类方数高达 32 首,占总方剂数量的 28.5%,所以炙甘草、桂枝、大枣等药物的出现频次要远大于其他药物,在最后划分结果中易被划分为独立节点,不利于其与其他药物的配伍规律研究。为减少高频药物对最终划分结果的影响,本文选用《伤寒论》中治疗少阴病证的 18 首方剂做了进一步分析,此 18 首方剂不包括桂枝汤及其类方且其方剂数量适中。

2.2 《伤寒论》少阴病证模型构建及社团划分

2.2.1 模型构建 《伤寒论》中少阴病证共包含四逆汤、通脉四逆汤等 18 首方剂,基于此 18 首方剂构建的“方-药”网络模型(简称“少阴病证模型”)如图 4 所示。其中节点的大小反映药物的药物影响力强弱、颜色代表不同的药物分类。图 4 共包含 36 个节点代表 18 首方剂中的 36 味中药,76 条边代表 18 首方剂中存在的 76 种“中药-中药”间关系。节点的大小反映出不同于《伤寒论》整体模型,少阴病证模型中药物影响力较大的 2 味药为炮附子和茯苓。结合少阴病证实际用药规律分析可发现,少阴

病为外感病发展过程中的危重阶段,病性多属虚、属阴、属寒,且少阴包括心肾二脏及所属经脉,而肾阳虚衰可致水汽泛滥,故其证治多用附子以补火助阳、莠寒除湿并佐以茯苓利水渗湿^[18]。

2.2.2 社团划分 经过后续参数实验(见“2.3.1”),本文最终选取 $\alpha=0.8$ 、 $\beta=0.7$ 、 $\gamma=0.2$ 、 $\theta=0.9$ 进行了少阴病证模型社区划分,划分结果共包含 5 个社区,将其用 Cytoscape 软件进行可视化显示得到的最终结果如图 5 所示。

图 5 共包含 27 味药物,不同的节点颜色代表划分后不同的社团分类,如图中的黄色药物均属于社团 1、红色药物均属社团 4,蓝色药物则代表划分后社区中的重叠药物节点。分析图内药物的分布情况可以看出划分后的少阴病证模型具有明显的社区结构。其中社区 1 包含生附子、干姜等 5 味药;社区 2 包含枳实与柴胡 2 味药;社区 3 包含药物滑石、泽泻、猪苓 3 味药;社区 4 包含麻黄、细辛等 4 味药;社区 5 包含黄芩、黄连等 6 味药;而芍药、炮附子等 7 味药则为社区间的重叠药物节点。

结合药物功效及少阴病证的实际用药特点分析可发现,社区 1 中的生附子、干姜、葱白合用有破阴回阳、宣通上下之效,配合人尿、猪胆汁使用则能达到补精血、增阴液的目的,适用于阴脱阳竭、寒热格拒之证,此 5 味药配合可用于阳气虚衰、阴寒内置的少阴寒化证^[1]。社区 2 中的柴胡辛散升阳、疏肝解郁,枳实苦泄沉降、下气消痞、理气除满,二者一升一降、一肝一脾,配伍使用具有升降气机、调理肝脾的功效,是治疗中气下陷的常用药对^[24]。社区 3 中的滑石、泽泻、猪苓,这 3 味药同属利水渗湿药,具有通利水道、渗泄水湿的效果,可用于治疗少阴病中肾阳虚衰导致的水汽泛滥等症^[25]。社区 4 中的麻黄性温质轻为发汗解表第一要药、细辛辛温能祛内寒而温脏腑,两药同用既能鼓动内外之阳气以加强解表散寒之力,又有温肺化饮、散寒平喘之效^[24];吴茱萸味辛苦而性热,既能温胃暖肝以驱寒,又善和胃降逆而止呕,与大枣同属吴茱萸汤的重要组分^[1],是治疗肝胃虚寒、浊阴上逆的常用药。社区 5 中的黄芩与黄连可泻心火,鸡子黄善养心滋肾,3 药合用可用于治疗肾阴亏虚、心火旺神之证^[1],半夏与苦酒合用可增强涤痰敛疮之效^[1],半夏与桂枝合用有通阳散寒、涤痰开结之效^[1],这 3 味药配合使用可用于邪郁少阴经脉所致的咽痛证。重叠节点的大小反映出在少阴病证模型中芍药、

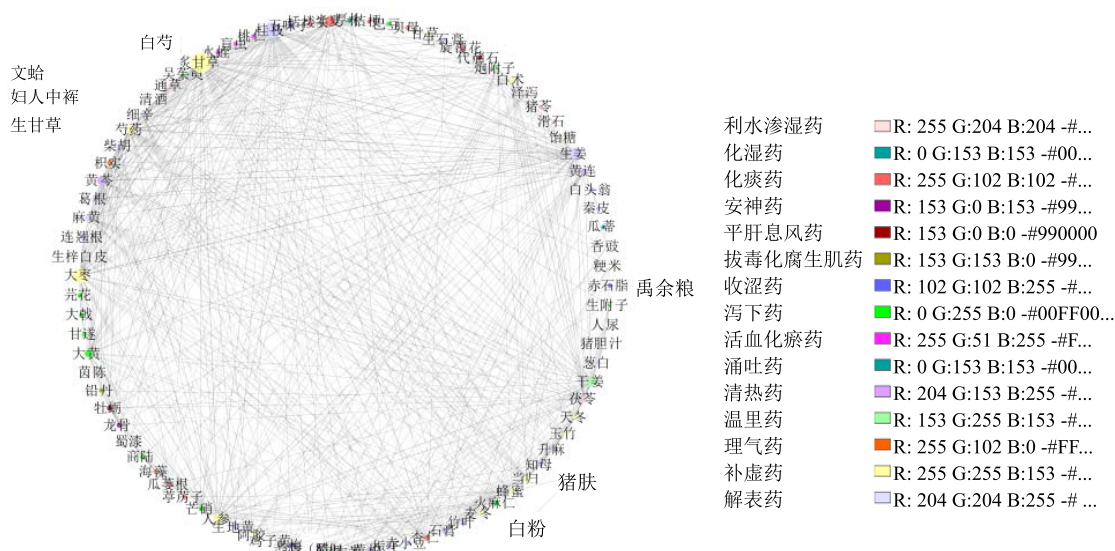


图 2 《伤寒论》整体模型

Fig. 2 Whole model of *Treatise on Cold Damage*

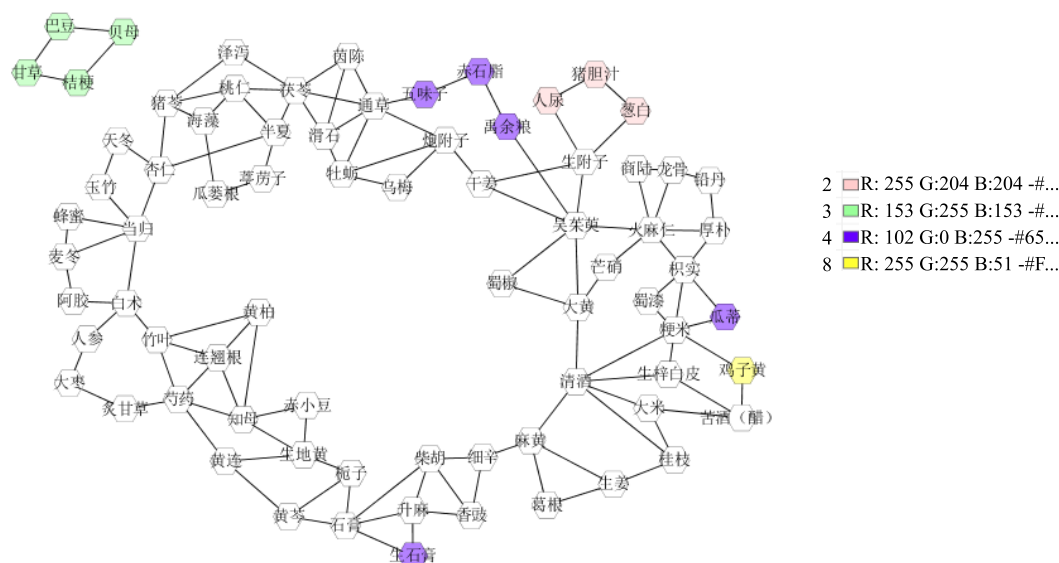


图 3 《伤寒论》整体模型社区划分结果

Fig. 3 Results of community division of whole model in *Treatise on Cold Damage*

炮附子、茯苓、阿胶、白术、生姜、人参这 7 味药的药物影响力均比较大, 其中芍药、炮附子、茯苓 3 味药的药物影响力在 36 味药物中排前 3, 而剩余阿胶、白术、生姜、人参 4 味药的药物影响力排前 10。较大的药物影响力反映出这 7 味药是治疗少阴病证的常用药物, 通过与其余药物配合使用共同发挥作用, 故在少阴病证模型社区划分过程中易被划分为重叠药物节点。

综上所述, 运用本文提出的“方-药”网络划分

方法进行《伤寒论》少阴病证模型社区划分时, 可得到结构清晰的药物社区, 且同社区中的药物可通过配伍发挥某种特定功效。

2.3 参数分析

2.3.1 “方-药”网络重叠社区划分方法参数 本文提出的适用于“方-药”网络的重叠社区划分方法主要涉及 4 个参数: α 、 β 、 γ 和 θ 。由于不同的数据集会得到不同的药物与药物间关系, 所以对于《伤寒论》整体模型和少阴病证模型的社区划分, α 、 β 、

表 2 重叠药物节点分类结果

类别	药物
1	黄芩、半夏
2	通草、吴茱萸、栀子、阿胶、火麻仁、麦冬、知母、粳米
3	黄连、柴胡、大黄
4	乌梅、黄柏、苦酒(醋)、大米、蜀椒、葛根、龙骨、铅丹
5	蜀漆、瓜蒌根、葶苈子、海藻、商陆、桃仁
6	杏仁、白术、炮附子
7	石膏、当归、细辛
8	生附子、清酒、赤小豆、香豉、生梓白皮、连翘根、茵陈、生地黄
9	猪苓、泽泻、滑石、天冬、玉竹、竹叶、升麻
10	枳实、蜂蜜、牡蛎、芒硝、厚朴
单独分类节点	大枣、炙甘草、生姜、桂枝、芍药、麻黄、茯苓、干姜、人参

表 3 部分药物出现频次统计

药物	频次	药物	频次
炙甘草	64	人参	22
桂枝	38	半夏	18
大枣	38	大黄	16
生姜	36	黄芩	16
芍药	30	麻黄	13
干姜	23	茯苓	13

γ 和 θ 的取值并不完全相同。但其核心思想均是通过比较核心药物社区规模、社区扩展规模以及最终得到的社区分类结果实现参数的选择。本文以《伤寒论》整体模型社区划分过程为例对相关参数值的实际选取进行了分析。

(1) 参数 β : 参数 β 的作用原理是通过控制药物

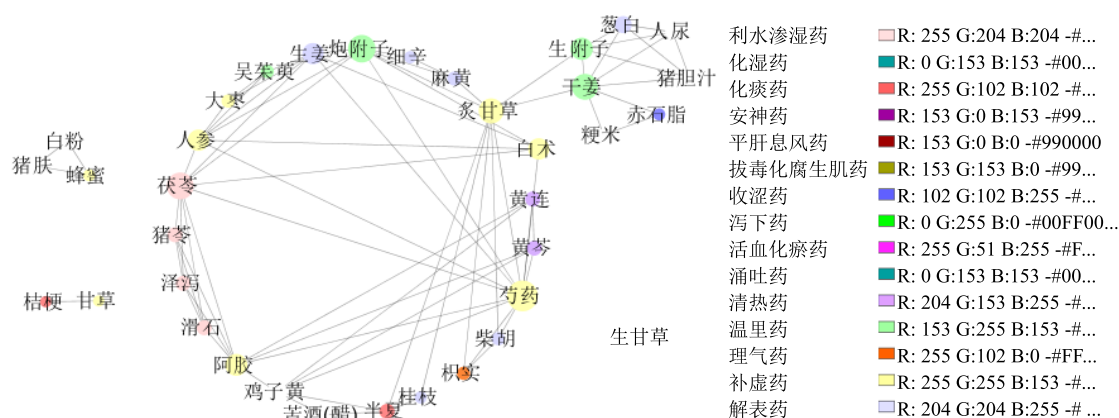


图 4 《伤寒论》少阴病证模型

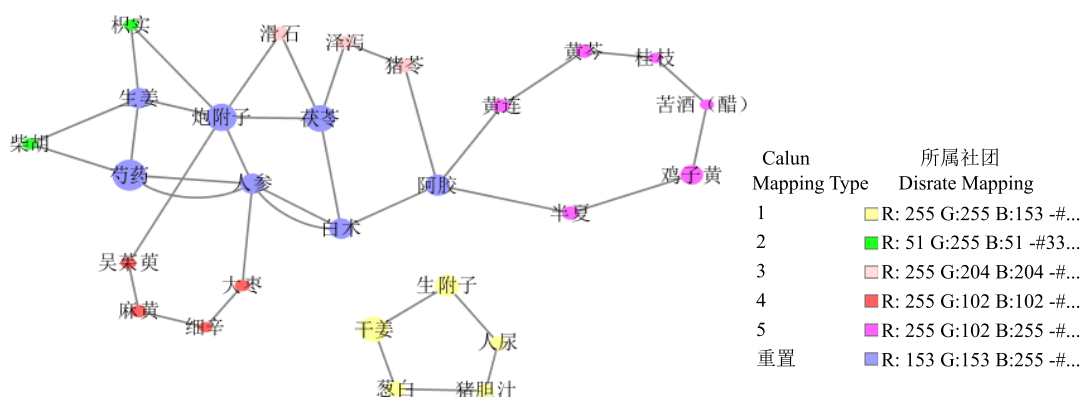
Fig. 4 Shaoyin syndrome model of *Treatise on Cold Damage*

图 5 《伤寒论》少阴病证模型社区划分结果

Fig. 5 Results of community division of Shaoyin syndrome model in *Treatise on Cold Damage*

影响力比当前药物小的邻节点数目与总邻节点数目的比值,来控制核心药物种子节点的个数,并与参数 γ 结合使用得到稳定紧密的核心药物社区。本文在保证参数 γ 一定的同时使用《伤寒论》整体数据集在 β 的范围 $[0, 1]$ 内进行了一系列实验,结果见图 6。

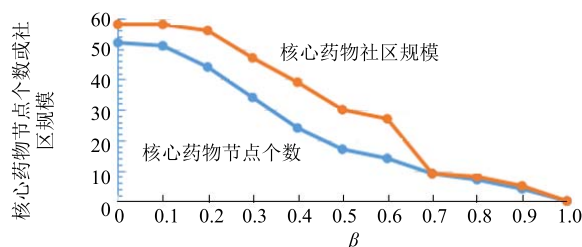


图 6 参数 β 的分析

Fig. 6 Analysis of parameter β

由图 6 显示,不同 β 值得到的核心药物节点个数不同且最终发现的社区规模也不同。根据 β 的作用原理及《伤寒论》整体模型的实际特点可知, β 越大,所识别核心药物的药物影响力越大,质量越高,但数量越少; β 越小,所识别核心药物的药物影响力越小,质量越低,但数量越多。为了保证药物节点质量并使用合理的药物节点扩展出较大的社区,分析不同 β 值下的核心药物节点个数和核心药物社区规模差值可发现,在《伤寒论》整体模型划分过程中参数 β 的取值范围应为 0.2~0.6。

(2) 参数 γ : 参数 γ 通过控制贡献度阈值选取社区邻节点的加入。本文选取 $\beta=0.4$ 时划分《伤寒论》整体模型得到的 24 个核心药物节点为 24 个初始社区,在 γ 的范围 $[0, 1]$ 内进行了一系列核心药物社区发现实验,结果如图 7 所示。

由图 7 显示,使用不同的 γ 值得到的核心药物社区规模不完全相同。而当 $\gamma>0.4$ 时,核心药物社区规模与原始社区规模相同,表明大于 0.4 的 γ 值对划分《伤寒论》整体模型不产生影响。为了进一步分析 γ 取值与社区规模间的关系,对 γ 值在 $[0, 0.4]$ 进行实验,结果如图 8 所示。

由图 8 显示,当 $\gamma>0.1$ 时核心药物社区规模和原始社区规模差值不大,表明 $\gamma>0.1$ 对于社区规模扩张影响较小。结合《伤寒论》整体模型的实际特点可发现,划分初期的社区中只包含少量核心药物,而这些药物在整个网络中同多种药物相连,所以这些药物对社区贡献的权值要远小于其在整个网络中的总权值,具体表现便是大多数药物的“药物对

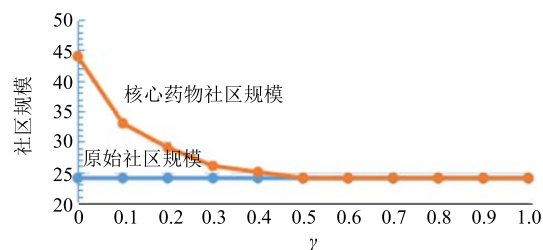


图 7 参数 γ 的分析 ($\gamma \in [0, 1]$)

Fig. 7 Analysis of parameter γ ($\gamma \in [0, 1]$)

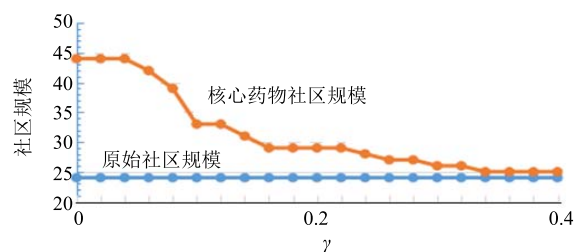


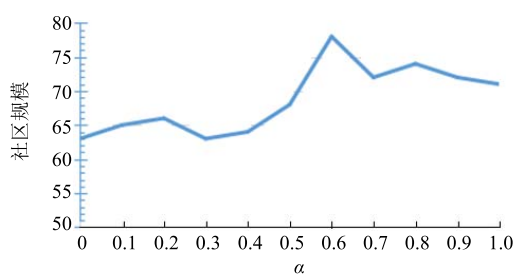
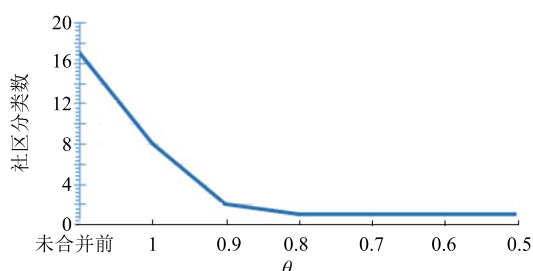
图 8 参数 γ 的分析 ($\gamma \in [0, 0.4]$)

Fig. 8 Analysis of parameter γ ($\gamma \in [0, 0.4]$)

社区贡献度”值均小于 0.1。为了保证 γ 的合理,在《伤寒论》整体模型划分实验中其实际取值范围应为 0~0.1。

(3) 参数 α : 参数 α 是本文定义的自适应函数中一个正的实数,其作用原理是通过控制 k_{in}^g 与 $(k_{in}^g + k_{out}^g)^\theta$ 的比值,实现社区的扩展。在保证参数 β 及 γ 一定的情况下,本文基于《伤寒论》整体模型在 α 取值范围 $[0, 1]$ 内进行了一系列实验,实验结果如图 9 所示。由图 9 显示,不同的 α 值所能发现的社区规模不同。由于《伤寒论》整体模型中药物影响力大的药物基本上与模型中的绝大多数药物存在连边,所以实际 k_{out}^g 的值要远大于 k_{in}^g 的值。故在本实验中 α 越大,单个社区越容易接纳更多的节点; α 越小,单个社区规模越小。但进一步分析实际划分后社团中包含的药味情况可发现,过大的 α 值得到的最终社区内包含着大量的重叠药物,为了使尽量多的药物在划分结果内至少属于一个社区,在《伤寒论》整体模型划分实验中参数 α 的取值范围应为 0.5~0.8。

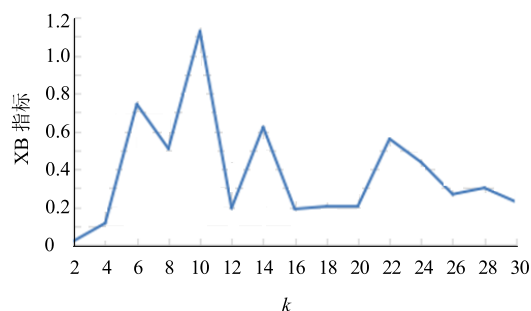
(4) 参数 θ : 参数 θ 通过控制社区相似度阈值来优化社区的合并。通常将相似度大于 0.5 的 2 个社区合并为同一个社区^[10],为了提高社区质量,本文基于《伤寒论》整体模型并在 θ 的范围 $[0.5, 1.0]$ 进行了一系列实验,结果如图 10 所示。

图 9 参数 α 的分析Fig. 9 Analysis of parameter α 图 10 参数 θ 的分析Fig. 10 Analysis of parameter θ

由图 10 显示, 当 $\theta < 0.9$ 时, 整个社区被重新合为一个社区, 表明《伤寒论》整体模型划分后的社区间存在着很高的重叠性。为了尽可能避免由于合并造成的社区过大而导致用药规律难以分析的问题, 在《伤寒论》整体模型划分实验中 θ 值范围应为 0.9~1.0。

2.3.2 FCM 算法参数 FCM 算法主要涉及 2 个参数, 即 m 和 k , 其中 m 为模糊加权指数, 用来控制聚类结果的模糊程度, k 为聚类个数。目前对 m 的取值已经有了一些研究, 但尚未得到有效进展, 同大多数 FCM 算法实验一样, 本文中 $m=2$ 。 m 的不同决定了最终聚类划分结果的好坏, XB 指标^[26]是目前常用的用于评价聚类好坏的指标, 本文在 k 的取值范围 [2, 30] 内结合 XB 指标进行了一系列实验。 k 值范围主要是依据聚类节点的个数确定, 本文包括 62 个药物节点, 为了尽可能的使同一类中包含 2 种及 2 种以上药物, k 值范围应为 [2, 30]。实验结果如图 11 所示。

由图 11 显示, 当 k 值不同时, 最终的 XB 指标值也不同。由 XB 指标的定义可知, XB 指标计算结果越小, 聚类效果越好。结合实际划分结果可得, 本研究中合理的 k 值范围为 [12, 20]。

图 11 参数 k 的分析Fig. 11 Analysis of parameter k

3 讨论

药物的功用各有所长、各有所短, 配伍是通过药物间的相互作用达到增效和减毒的作用^[25]。方剂是依据病情在辨证立法的基础上遵循“君、臣、佐、使”的基本组织结构, 选择合适的药物配伍而成^[1], 内涵丰富的复杂性规律。复杂网络作为分析复杂现象的一种常用方法, 近年来已被应用于方剂的知识发现分析, 如洪文学团队将复杂网络与中医数量化相结合, 借用复杂概念网络生成方法实现了方剂、药物和“方剂-药物”之间关系结构的可视化表示^[27], 并在中医辨证^[28]、方剂配伍^[29]、《伤寒论》知识发现^[27]等研究过程中取得了诸多成果。Ma 等^[30]借鉴 SNA (social network analysis) 思想并结合中医病案的特点对中药数据进行分析, 运用改进的度中心性、紧密中心性和介数中心性 3 种分析算法提取了中医病案中的核心药物。李梢等^[31]从网络的角度建立了一种基于距离的互信息模型 (distance-based mutual information model), 通过对药物相互作用进行评分构建药物网络。Gao 等^[32]构建了一种基于 Jaccard 相似系数的中药方剂网络。孙道平等^[33]提出了一种用于中药方剂网络重叠社区发现的改进 COPRA 算法。孙正^[34]依据药物间关联强度构建了中药方剂网络, 并提出了一种结合网络结构和节点属性的重叠社团发现算法。

本研究首先借助复杂网络的思想以方剂中药物为节点, 药间关系为边、药物共现频次表示边的权重, 构建了“方-药”网络模型, 在模型构建过程中引入“药物影响力”的概念识别模型中的重要药物节点。基于《伤寒论》112 首方剂构建的“方-药”网络模型显示, “药物影响力”大的 5 味药为炙甘草、桂枝、大枣、生姜、半夏, 这 5 味药物符合张仲景治疗伤寒疾病时“寒温并用、辛甘苦同行”的用药

原则, 其中前 4 位药合用可用于外感风寒、营卫不和所致的恶风发热然有汗出的表虚证; 半夏配生姜降逆止呕可用于治痰饮呕吐、半夏配桂枝宣散郁阳临床上可用于治疗面神经麻痹; 基于《伤寒论》少阴病证中 18 首方剂构建的“方-药”网络模型也显示, “药物影响力”较大的炮附子和茯苓 2 味药物符合少阴病证“温经扶阳、利水渗湿”的用药特点, 两药配伍有温阳利水宁心之效, 可用于治疗阳虚寒湿内侵之证, 上述模型构建结果表明, 本研究提出的“药物影响力”概念可用于“方-药”网络中重要药物节点的识别。其次, 提出了一种基于重叠社区的“方-药”网络划分方法, 该方法借鉴 i-SEOCDA 算法^[10]并结合实际构建的加权“方-药”网络模型, 通过引入“药物对社区相似度”“自适应函数”“社区相似度”等概念实现了“方-药”网络中重叠社区的发现。最后, 运用提出的“方-药”网络重叠社区划分方法对构建的《伤寒论》整体模型及少阴病证模型进行了划分。在《伤寒论》整体模型划分过程中, 引入了 FCM 算法^[16]对划分后的重叠药物节点进行分类, 分类结果显示“方-药”网络划分方法结合 FCM 算法, 可发现隐藏在《伤寒论》整体模型中的常用药组。如有通腑泄热、疏肝利胆作用的大黄及柴胡; 可用于熄风通络的生石膏及细辛; 可治气机郁结于胃肠之证的枳实及厚朴等。而基于《伤寒论》少阴病证模型的社区划分结果显示, 运用本文提出的方法可得到功效不同的药物配伍社区。如同属于一个社区的生附子、干姜、葱白、人尿、猪胆汁 5 味药相互配伍, 可用于治疗阳气虚衰、阴寒内盛的少阴寒化证; 同属于一个社区的黄芩、黄连、鸡子黄 3 味药合用, 可用于治疗肾阴亏虚、心火旺神之证, 而半夏、苦酒与桂枝 3 味药配合使用, 可用于邪郁少阴经脉所致的咽痛证。以上划分结果均证实了“方-药”网络中存在高质量的重叠社区结构, 运用本文划分方法可发现其中的社区, 并可依据划分结果进行相关用药规律的研究。

然而“方-药”网络模型在实际构建过程中未考虑药物剂量对组方的影响, 且划分后得到的药物分类结果并未囊括全部药物, 存在一定的局限。未来的工作将考虑在模型构建过程中加入“药物计量”“君臣佐使”“证型”等指标, 实现“理-法-方-药”的综合考量, 为阐明复方配伍的科学内涵、完善中药药性理论、指导中药复方开发等提供依据。

参考文献

- [1] 李 冀. 方剂学 [M]. 第 10 版. 北京: 中国中医药出版社, 2016.
- [2] 左明晏, 许从莲. 方剂配伍规律的研究概况 [J]. 中医杂志, 2016, 57(3): 260-263.
- [3] 邢雁辉, 崔 蒙, 储戟农, 等. 基于贝叶斯分类算法的治疗中风中药组方研究 [J]. 中西医结合心脑血管病杂志, 2015, 13(4): 471-474.
- [4] 刘 广, 孙艳秋. 基于 K-Means 聚类算法的消渴方剂研究 [J]. 中华中医药学刊, 2017, 35(1): 173-178.
- [5] 韩 楠, 乔少杰, 宫兴伟, 等. 面向正负关联规则的方剂配伍规律挖掘算法 [J]. 小型微型计算机系统, 2017, 38(7): 1538-1542.
- [6] 吕庆莉. 数据挖掘与复杂网络的融合及其在中医药领域应用 [J]. 中草药, 2016, 47(8): 1430-1436.
- [7] Fortunato S. Community detection in graphs [J]. *Physics Rep*, 2010, 486(3/5): 75-174.
- [8] Atzmueller M, Doerfel S, Mitzlaff F. Description-oriented community detection using exhaustive subgroup discovery [J]. *Inform Sci*, 2016, doi: 10.1016/j.ins.2015.05.008.
- [9] Lancichinetti A, Fortunato S, Kertesz J. Detecting the overlapping and hierarchical community structure in complex networks [J]. *New J Physic*, 2009, doi: 10.1088/1367-2630/11/3/033015.
- [10] 於志勇, 陈基杰, 郭 昆, 等. 基于影响力与种子扩展的重叠社区发现 [J]. 电子学报, 2019, 47(1): 153-160.
- [11] Su Y S, Wang B J, Zhang X Y. A seed-expanding method based on random walks for community detection in networks with ambiguous community structures [J]. *Sci Rep*, 2017, doi: 10.1038/srep41830.
- [12] Whang J J, Gleich D F, Dhillon I S. Overlapping community detection using neighborhood-inflated seed expansion [J]. *IEEE Transact Knowledge Data Eng*, 2016, 28(5): 1272-1284.
- [13] Chen D, Fu Y, Shang M. An efficient algorithm for overlapping community detection in complex networks [A] // 2009 WRI Global Congress on Intelligent Systems [C]. Xiamen: IEEE, 2009.
- [14] Gregory S. Finding overlapping communities in networks by label propagation [J]. *New J Physic*, 2010, doi: 10.1088/1367-2630/12/10/103018.
- [15] 刘永瑞, 于 晶, 刘 南. 基于数据挖掘的张仲景《伤寒论》用药规律探讨 [J]. 亚太传统医药, 2019, 15(7): 177-179.
- [16] Bezdek J C, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm [J]. *Comp Geosci*, 1984, 10(2/3): 191-203.

- [17] 赵立昌. 经方药对浅析 [J]. 河南中医, 2019, 39(8): 1158-1161.
- [18] 刘俊红. 治疗急腹症方剂中常用药对 [J]. 辽宁中医药大学学报, 2013, 15(4): 145-147.
- [19] 王 付. 仲景方黄连药对研究与应用 [J]. 中医药通报, 2007, 6(5): 28-29.
- [20] 李 云. 细辛药对的配伍效用分析 [J]. 黑龙江中医药, 2016, 45(4): 59-60.
- [21] 王伟杰, 唐晓娟. 浅谈细辛在风湿病中的应用 [J]. 上海中医药杂志, 2016, 50(8): 67-69.
- [22] 潘 玥, 陆为民. 徐景藩运用含厚朴药对治疗脾胃病经验 [J]. 中国中医药图书情报杂志, 2019, 43(2): 61-63.
- [23] 任 园, 麻春杰. 基于中医药传承辅助平台探讨《伤寒论》和《金匱要略》的用药规律 [J]. 江苏中医药, 2019, 51(9): 73-75.
- [24] 胥庆华. 中药药对大全 [M]. 北京: 中国中医药出版社, 1996.
- [25] 高学敏. 中药学 [M]. 北京: 中国中医药出版社, 2002.
- [26] Xie X L, Beni G. A validity measure for fuzzy clustering [J]. *Trans PAMI*, 1991, 13(4): 841-847.
- [27] 刘超男. 基于形式概念分析数学理论的《伤寒论》方药知识发现 [D]. 广州: 广州中医药大学, 2013.
- [28] 梁嘉鸣. 基于属性偏序原理探讨《伤寒杂病论》呕吐证治规律 [D]. 广州: 广州中医药大学, 2017.
- [29] 刘树林, 徐笋晶, 李日东, 等. 基于属性偏序原理的仲景经方咳嗽治疗规律知识发现 [J]. 燕山大学学报, 2014, 38(5): 455-459.
- [30] Ma Y K, Zhang D Z, Wulamu A, *et al.* The core drugs analysis based on social network analysis about traditional Chinese medicine records semantic relation [J]. *Proced Comp Sci*, 2014, doi: 10.1016/j.procs.2014.05.275.
- [31] Li S, Zhang B, Jiang D, *et al.* Herb network construction and co-module analysis for uncovering the combination rule of traditional Chinese herbal formulae [J]. *BMC Bioinform*, 2010, 11(11): S6.
- [32] Gao Y, Zheng L, Wang C J, *et al.* Chinese medicine formula network analysis for core herbal discovery [A] // International Conference on Brain Informatics [C]. Heidelberg: Springer, 2012.
- [33] 孙道平, 高原, 谢 隽, 等. 一种用于中药方剂网络重叠社区发现的改进 COPRA 算法 [J]. 南京大学学报: 自然科学版, 2013, 49(4): 483-490.
- [34] 孙 正. 基于药物属性的中药方剂组网及药物社团发现研究 [D]. 南京: 南京大学, 2013.