

信息技术应用

基于最小路径覆盖算法的化工过程重要参数的识别

秦艳¹, 徐一凡¹, 杨燕霞¹, 王政^{1*}, 贾小平², 王芳²

(1. 青岛科技大学化工学院, 山东 青岛 266042;

2. 青岛科技大学环境与安全工程学院, 山东 青岛 266042)

摘要:将复杂网络的目标控制理论应用到化工过程系统重要参数的识别中,以 SDG (signed directed graph) 模型和复杂网络理论为基础构建化工过程的网络模型,然后利用 LeaderRank 和节点相似度算法 (SRank 算法) 对网络节点重要性进行排序并基于此对网络进行鲁棒性分析选取目标节点,通过最小路径覆盖算法对网络进行目标控制分析,确定驱动节点并对它们进行重点监控。案例分析结果表明,该方法可行,对化工过程系统中重要参数的监测和安全控制具有一定的指导意义。

关键词:化工过程;复杂网络;目标控制;SRank 算法;鲁棒性;最小路径覆盖算法

中图分类号: X92

文献标志码: A

文章编号: 0253-4320(2019)04-0202-05

DOI: 10.16606/j.cnki.issn.0253-4320.2019.04.047

Identification of important parameters in chemical process based on minimum path coverage algorithm

QIN Yan¹, XU Yi-fan¹, YANG Yan-xia¹, WANG Zheng^{1*}, JIA Xiao-ping², WANG Fang²

(1. College of Chemical Engineering, Qingdao University of Science and Technology, Qingdao 266042, China;

2. College of Environment and Safety Engineering, Qingdao University of Science and Technology,

Qingdao 266042, China)

Abstract: This paper proposes the application of complex network target control theory in the identification of important parameters in chemical process system. The network model of chemical process is built on the basis of SDG (signed directed graph) model and complex network theory, then the importance of network nodes is sorted by LeaderRank and node similarity algorithm (SRank algorithm). Based on this, the robustness of the network is analyzed and the target node is selected. The target control analysis is carried out by the minimum path coverage algorithm and the driving nodes are confirmed and monitored. Case studies show that this method is feasible and has some guiding significance for monitoring and safety control of important parameters in chemical process system.

Key words: chemical processes; complex network; target control; SRank algorithm; robustness; minimum path coverage algorithm

随着化工生产规模的不断扩大,识别化工过程的重要参数并对其进行有效地监测和控制,能够降低化工生产过程的运行风险,对化工生产过程的安全稳定运行具有重要意义。目前,利用复杂网络理论与 SDG 模型来识别化工过程系统重要节点的研究越来越受到广泛关注。王政等^[1]利用 SDG 模型将化工过程转化为网络模型,通过统计特征描述和社团划分确定了网络的核心节点;Cai 等^[2]将复杂网络理论和主成分分析方法相结合对化工过程

进行监控并检测出系统故障;姜英等^[3]利用层次分析法构建化工过程层次 SDG 模型并转化为网络结构,识别出网络的关键节点等。这些方法多是基于网络拓扑结构和统计特征来识别重要节点,但未从复杂网络控制理论的角度考虑驱动节点的重要性。

通常情况下,控制整个网络所需要的成本较高,复杂网络目标控制是对于给定的网络找到最小的驱动节点集合并加入输入信号使其能够控制所有的目

收稿日期:2018-09-18;修回日期:2019-01-19

基金项目:国家自然科学基金项目(41771575)

作者简介:秦艳(1992-),女,硕士生;王政(1968-),男,博士,副教授,研究方向为过程系统工程,通讯联系人,wangzheng@qust.edu.cn。

标节点,使网络能够安全稳定地运行^[4],这样既可以降低生产成本又能够有效地控制目标节点。复杂网络目标控制的研究已经取得了一定的成果,例如,Gao等^[4]提出了贪婪算法求解网络目标控制所需的驱动节点,但其结果受到最大匹配集选取的影响;Zhang等^[5]在贪婪算法的基础上提出优先匹配原则,降低了驱动节点数但不能保证是最优结果;薛等长^[6]提出的最小路径覆盖算法得到的结果不会受到最大匹配集选取的影响且具备更低的计算复杂度等。复杂网络目标控制已经应用于生物学^[7-8]、癌症诊断^[9]、经济网络^[10]等领域。随着应用复杂网络研究化工过程系统的深入,化工过程网络鲁棒性研究的理论意义与应用价值也日益凸显。网络鲁棒性是指网络在外界蓄意攻击或自身随机故障的情况下,网络系统还能维持自身功能的能力^[11]。目前,网络鲁棒性的研究已应用到公交网络^[12]、电力网络^[13]、供应链网络^[14]等领域。对化工过程网络中的节点进行蓄意攻击,根据评价指标确定网络接近崩溃状态时网络的失效节点并以此作为目标节点,有利于发觉化工过程网络的重要节点并进行有效保护,保证化工生产过程的安全运行。

因此,本文中提出将复杂网络的目标控制理论应用到化工过程系统的重要参数识别。以流程图建模法构建化工过程SDG模型并转化为有向网络模型,然后采用SRank算法对网络节点重要性进行评估,根据节点重要性排序对网络进行蓄意攻击分析其鲁棒性并以此确定目标节点,通过最小路径覆盖算法对网络进行目标控制分析,确定网络的驱动节点集并对其进行重点监控,从而可以更加准确和全面地辨别网络中需要重点监控的化工过程参数,降低化工生产过程的运行风险。

1 网络模型的建立

SDG模型通过用节点表示系统中的状态变量,有向边表示变量间的影响关系,从而将化工过程抽象成符号有向图模型^[15]。本文中将SDG模型中的变量看作网络的节点,它们之间的关系抽象为网络的边,构建邻接矩阵 A ,其中的元素 a_{ij} 代表节点 i 与节点 j 之间的关联度,见式(1):

$$a_{ij} = \begin{cases} 1 & \text{节点 } i \text{ 指向节点 } j \\ 0 & \text{节点 } i \text{ 不指向节点 } j \end{cases} \quad (1)$$

利用Ucinet软件输入邻接矩阵 A ,通过Netdraw功能即可得到相应的有向网络模型。

2 基于复杂网络目标控制理论的化工过程重要参数的识别

2.1 基于SRank算法的节点重要性评价

综合考虑到节点的局部特点和全局特点以及网络的有向性,本文中采用SRank算法对网络节点进行重要性排序。顾亦然等^[16]提出了SRank节点重要性排序算法,即在LeaderRank算法基础上考虑节点相似度。

2.1.1 LeaderRank算法

(1)初始时刻赋予网络中除了背景节点以外的所有节点一个单位值,即 $LR_i(0) = 1, LR_g(0) = 1$ 。

(2) $LR_i(k)$ 为经过 k 步迭代后系统到达稳态,则节点 i 的初步重要度计算如下:

$$LR_i(k) = \sum_{j=1}^{n+1} (a_{ij}/k_j^{out}) LR_j(k-1) \quad k = 1, 2, 3, \dots \quad (2)$$

式中, n 为去除背景节点的网络节点数;若节点 i 指向节点 j ,则 $a_{ij} = 1$,否则为0; k_j^{out} 为节点 j 的出度。

(3)达到稳态后,背景节点将自身影响力均匀分配到网络中的其他各个节点,则可以得到网络中节点的全局重要度,见式(3),其中 $LR_g(k)$ 为第 k 步背景节点的重要度。

$$LR_i = LR_i(k) + LR_g(k)/n \quad (3)$$

2.1.2 节点相似度

节点 i 和它的邻居节点 j 的相似度 Sim_{ij} 的定义见式(4):

$$Sim_{ij} = \gamma k_{ij}^{out} + (1 - \gamma) k_{ij}^{in} \quad j \in \alpha_i \quad (4)$$

$$SR_i = LR_i \sum_{j \in \alpha_i} [(Sim_{ij} + 1)/(k_i^{out} + 1)] LR_j \quad (5)$$

式中, α_i 为节点 i 的邻居节点的集合; γ 为调节参数,表示节点出度和入度的相对重要性,取值范围为 $0 < \gamma < 1$; k_{ij}^{out} 表示节点 i 和 j 同时指向某一节点的节点个数; k_{ij}^{in} 表示同时指向节点 i 和 j 的节点个数。

基于SRank算法的节点重要性排序步骤:首先,用LeaderRank算法计算网络中节点的 LR 值;其次,根据式(4)计算节点对之间的相似度;最后,根据式(5)计算节点的 SR 值得节点重要性排序。

2.2 蓄意攻击下的网络鲁棒性分析

2.2.1 评价指标

(1)最大连通度

当网络遇到攻击时,最大连通度可以较好地描述网络中各节点的连通程度。最大连通度 $S(G)$ 是指当网络受到攻击或干扰时,网络中的最大子网络中所含的节点数量与网络中所有节点数量的比值,

见式(6):

$$S(G) = n/N \quad (6)$$

式中, n 指最大连通子图的节点数; N 指网络总节点数。

$S(G)$ 越大, 网络连通性越高, 则鲁棒性越强。

(2) 网络效率

当网络中的节点受到一定程度的攻击时, 会产生多个孤立节点, 利用 $S(G)$ 很难科学有效地显示出网络连通性的变化情况。本文将网络效率测度用来描述化工系统网络的鲁棒性。网络效率是指网络中所有节点对之间距离倒数之和的平均值, 见式(7):

$$S(G) = [1/N(N-1)] \sum_{i,j \in v, i \neq j} (1/d_{ij}) \quad (7)$$

式中, N 为网络节点的数目; d_{ij} 为节点 i 与节点 j 之间的距离。

2.2.2 攻击方式

采用蓄意攻击的方式对网络进行攻击, 对于网络的蓄意攻击方式, 大多数研究都是基于节点度和节点介数的攻击方式。由于节点介数和度对节点的评价具有片面性, 因此提出基于节点重要性排序的攻击方式, 根据节点重要性的大小依次使网络中的节点失效, 进而对其进行网络的鲁棒性分析。当网络在节点失效的情况下, 网络接近于崩溃状态时网络中的失效节点则为目标节点。

2.3 基于最小路径覆盖算法的化工过程网络目标控制

基于最小路径覆盖算法的具体步骤如下。

第 1 步: 输入网络的邻接矩阵, 同时根据节点重要性排序选取目标节点集 S_i , 并将所选的目标节点按照出度从大到小进行排序, 且将最小路径覆盖集 S_{path} 初始化为空集。

第 2 步: 依次计算出目标节点集 S_i 中从每个节点开始并终止于 S_i 终节点的路径, 并找到含有目标节点最多的一条路径, 记为 P_{max} 。若路径中所包含的目标节点数目相同, 则随机选择一条路径。

第 3 步: 判断 P_{max} 与 S_{path} 中的路径是否相交, 若不相交, 则将 P_{max} 加入 S_{path} 中, 并且将 P_{max} 中出现的所有目标节点从 S_i 中移除; 若相交, 则算法继续。

第 4 步: 如果目标节点集 S_i 为空集, 则算法停止。这时, 将最小路径覆盖集 S_{path} 中每一条路径的第一个节点作为目标控制的驱动节点; 否则, 转至第 2 步。

3 案例分析

3.1 案例一

3.1.1 网络模型的建立

以合成氨工艺为例进行案例分析, 流程图及相关变量的定义见文献[1]。通过各变量之间的关系建立合成氨工艺的 SDG 模型, 利用 Ucinet 软件构建出合成氨工艺的有向网络模型, 如图 1 所示。

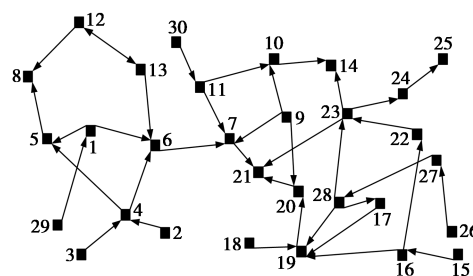


图 1 合成氨工艺的网络模型

3.1.2 基于 SRank 算法计算网络节点重要性排序

利用 MATLAB 计算节点的 SR 值得到节点重要性排序, 如表 1 所示。

表 1 合成氨网络节点的 SR 值及排序

排序	节点	SR 值	排序	节点	SR 值	排序	节点	SR 值
1	21	8.788	11	4	1.319	21	27	0.723
2	19	3.727	12	10	1.234	22	13	0.626
3	7	3.520	13	28	1.208	23	12	0.617
4	23	2.873	14	14	1.203	24	18	0.506
5	20	2.832	15	25	1.119	25	2	0.365
6	8	2.472	16	22	1.008	26	3	0.365
7	6	2.139	17	1	0.896	27	15	0.304
8	5	1.751	18	11	0.896	28	26	0.304
9	17	1.537	19	16	0.878	29	29	0.304
10	24	1.323	20	9	0.865	30	30	0.304

3.1.3 网络目标节点的选取

根据表 1 中节点重要性排序依次选择一定比例的节点进行失效并计算节点失效后网络的 $S(G)$ 和 $E(G)$, 将所得结果进行整理并作图, 如图 2 所示。网络在蓄意攻击下的鲁棒性较差, 节点失效比例达到 40% 时, 网络最大连通度低于 0.1, 网络效率低于 0.01, 网络几乎达到崩溃状态, 因此选择节点重要性在前 40% 的节点作为目标节点进行目标控制分析。

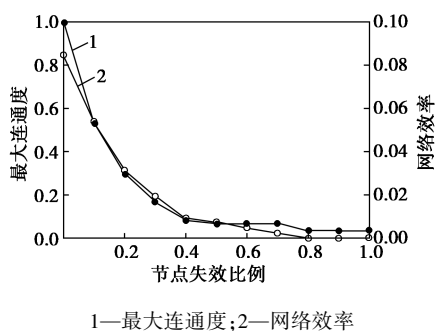


图2 合成氨网络 $S(G)$ 和 $E(G)$ 随节点失效比例的变化

3.1.4 基于最小路径覆盖算法的网络驱动节点

合成氨工艺网络的目标节点集为: $S_t = \{21, 19, 7, 23, 20, 8, 6, 5, 17, 24, 4, 10\}$ 。根据最小路径覆盖算法得到合成氨工艺网络的相关路径和驱动节点, 如表2所示。因此, 得到合成氨工艺网络的目标控制驱动节点集为 $S_d = \{17, 4, 23, 6, 10\}$, 即吸收塔液位、蒸汽流量、合成塔压力、一段转化炉流量和二段转化炉压力。这些变量比较重要, 如果发生故障会对网络中的其他节点造成一定的影响, 进而影响到整个化工生产系统的运行, 因此需要对其进行重点监控, 以确保生产过程安全稳定地进行, 从而预防事故的发生。该方法得到的结果相比于文献[1]中得到的结果更加详细。

表2 合成氨网络的目标控制驱动节点

路径	17→19→20→21	4→5→8	23→24	6→7	10
驱动节点	17	4	23	6	10

3.2 案例二

3.2.1 网络模型的建立

以原油蒸馏中的常压蒸馏为例进行案例分析, 工艺流程图及变量定义见文献[17]。分析各变量之间的关系建立该过程的SDG模型并构建邻接矩阵, 利用Ucinet软件得到该过程有向网络模型, 如图3所示。

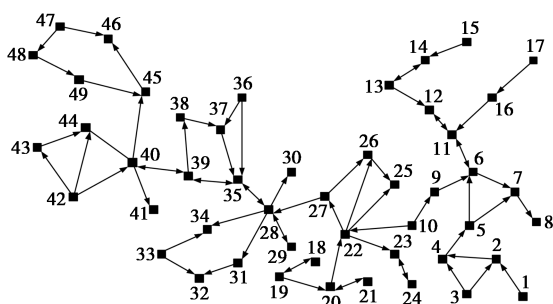


图3 常压蒸馏过程的网络模型

3.2.2 基于SRank算法计算网络节点重要性排序

利用MATLAB计算节点的SR值得到节点重要性排序, 如表3所示。

表3 常压蒸馏过程网络节点的SR值及排序

排序	节点	SR值	排序	节点	SR值	排序	节点	SR值
1	35	3.452	18	22	1.473	35	36	0.671
2	12	2.409	19	28	1.438	36	24	0.665
3	40	2.386	20	5	1.391	37	42	0.657
4	46	2.267	21	30	1.313	38	29	0.656
5	7	2.262	22	23	1.307	39	19	0.646
6	6	2.261	23	8	1.285	40	48	0.618
7	11	2.248	24	41	1.192	41	3	0.568
8	26	2.225	25	20	1.169	42	21	0.565
9	34	2.158	26	31	1.135	43	33	0.467
10	37	2.131	27	16	1.064	44	47	0.450
11	45	2.037	28	38	1.055	45	10	0.417
12	44	1.985	29	13	0.952	46	18	0.413
13	25	1.775	30	43	0.934	47	15	0.363
14	4	1.725	31	14	0.924	48	1	0.322
15	32	1.677	32	27	0.903	49	17	0.284
16	2	1.586	33	49	0.889			
17	39	1.509	34	9	0.712			

3.2.3 网络目标节点的选取

根据表3中节点重要性排序依次选择一定比例的节点进行失效并计算节点失效后网络的 $S(G)$ 和 $E(G)$, 将所得结果进行整理并作图, 如图4所示。网络在蓄意攻击下的鲁棒性较差, 当节点失效比例达到40%时, 网络最大连通度低于0.1, 网络效率接近于0.01, 网络几乎达到崩溃状态, 因此选择节点重要性在前40%的节点作为目标节点进行目标控制分析。

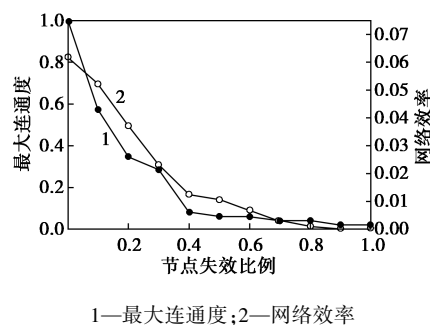


图4 常压蒸馏网络 $S(G)$ 和 $E(G)$ 随节点失效比例的变化

3.2.4 基于最小路径覆盖算法的网络驱动节点

常压蒸馏过程网络的目标节点集为: $S_1 = \{35, 12, 40, 46, 7, 6, 11, 26, 34, 37, 45, 44, 25, 4, 32, 2, 39, 22, 28, 5\}$ 。根据最小路径覆盖算法得到网络的相关路径和驱动节点,如表 4 所示。因此,得到常压蒸馏过程网络的目标控制驱动节点集为 $S_d = \{22, 2, 26, 32, 34, 37, 44\}$,即炉膛温度、电盐罐油水界面、炉膛温度调节阀、常压塔塔顶回流罐液位、常压塔塔顶压力、常一中进料温度和常二中回流温度。这些变量在网络中处于重要位置,如果发生故障会对网络中的其他变量造成一定的影响,进而影响到整个化工生产系统的运行,因此需要对其进行重点保护和监控,以确保生产过程安全稳定地进行,从而预防事故的发生。

表 4 常压蒸馏过程的目标控制驱动节点

路径	驱动节点
22→27→28→35→39→40→45→46	22
2→4→5→6→11→22	2
26→25	26
32	32
34	34
37	37
44	44

4 结论

(1) 在 SDG 模型和复杂网络理论的基础上,建立化工过程的有向网络模型,在节点重要性排序的基础上对网络进行鲁棒性分析并确定目标节点,通过最小路径覆盖算法对网络进行目标控制分析确定网络的驱动节点,进而获得需要重点监控的参数,从而为化工生产过程的安全运行、监测控制提供一定的理论依据,案例分析表明了本文中所提方法的可行性。

(2) 将复杂网络的目标控制理论应用到化工过程系统重要参数的识别,在节点重要性排序的基础上考虑网络目标控制中驱动节点的重要性,从而使化工过程中的重要参数识别更加详细和准确,为化工过程参数的监测和控制提供了新的解决思路。

参考文献

- [1] 王政,孙锦程,王迎春,等.基于复杂网络理论的符号有向图(SDG)化工故障诊断[J].化工进展,2016,35(5):1344-1352.
- [2] Cai E, Dan L, Ling L, *et al.* Monitoring of chemical industrial processes using integrated complex network theory with PCA[J]. Chemometrics & Intelligent Laboratory Systems, 2015, 140(27): 22-35.
- [3] 姜英,王政,秦艳,等.基于复杂网络的化工过程层次符号有向图模型建立及关键节点识别[J].化工进展,2018,37(2):444-451.
- [4] Gao J, Liu Y Y, D'souza R M. Target control of complex networks [J]. Nature Communications, 2014, 5: 5415-5423.
- [5] Zhang X, Wang H, Lv T. Efficient target control of complex networks based on preferential matching [J]. Plos One, 2017, 12(4): e0175375.
- [6] 薛等长.基于最小路径覆盖的复杂网络目标控制的研究[D].西安:西安电子科技大学,2015.
- [7] Sharma A, Cinti C, Capobianco E. Multitype network-guided target controllability in phenotypically characterized osteosarcoma: Role of tumor microenvironment [J]. Frontiers in Immunology, 2017, 8: 918-936.
- [8] Zhang X F, Ou-Yang L, Zhu Y, *et al.* Determining minimum set of driver nodes in protein-protein interaction networks [J]. Bmc Bioinformatics, 2015, 16(1): 146-159.
- [9] Kanhaiya K, Czeizler E, Gratie C, *et al.* Controlling directed protein interaction networks in cancer [J]. Sci Rep, 2017, 7(1): 1155-1167.
- [10] Galbati, Marco, Delpini, *et al.* The power to control [J]. Nature Physics, 2013, 9(3): 126-128.
- [11] 陆靖桥,傅秀芬,蒙在桥.复杂网络的鲁棒性与中心性指标的研究[J].计算机应用与软件,2016,33(4):302-305.
- [12] 鲍登,高超,张自力.基于复杂网络的公交-地铁复合网络鲁棒性分析[J].西南师范大学学报,2017,42(5):22-27.
- [13] 罗筱如.基于复杂网络理论的电力网络鲁棒性及脆弱性分析[D].成都:西南交通大学,2012.
- [14] 徐翔斌,李恒.基于供需能力的供应链网络鲁棒性[J].系统工程,2015,(8):17-23.
- [15] 张贝克,许欣,高东,等.基于定性趋势与符号有向图的模型校核方法[J].化工学报,2013,64(12):4536-4543.
- [16] 顾亦然,朱梓嫣.基于 LeaderRank 和节点相似度的复杂网络重要节点排序算法[J].电子科技大学学报,2017,46(2):441-448.
- [17] 姜英.基于复杂网络的化工过程层次 SDG 故障传播特性的研究[D].青岛:青岛科技大学,2018. ■

《现代化工》欢迎广大作者踊跃投稿,投稿系统: <http://www.xdhg.com.cn>