

一种基于网络最大可控子图的导航搜索模型*

肖延东[†] 老松杨 侯绿林 白亮

(国防科技大学信息系统与管理学院, 长沙 410073)

(2013年5月6日收到; 2013年9月24日收到修改稿)

基于网络可控性模型提出了最大可控子图的概念, 在此基础上提出了一种基于最大可控子图的导航搜索模型. 模型中基于最大可控子图的加边策略用最小的代价解决了有向网络搜索中存在的粒子因“无路可走”而终止搜索的问题; 基于最大可控子图部署导航节点, 仅用节点总数 2% 左右的导航点, 就使全网搜索时间接近导航网络的平均最短路径. 通过在 ER 和 SF 网络上的实验表明, 全网搜索时间与网络的可控性有关, 可控性越好, 添加的边数量越少, 同时会使网络中导航节点分布越多, 越能提高网络的搜索效率.

关键词: 导航搜索, 有向网, 网络可控性

PACS: 89.75.Hc, 89.75.Fb, 02.30.Yy

DOI: 10.7498/aps.62.248901

1 引言

近年来, 随着互联网、社会关系网、基因调控网以及新陈代谢网等大规模复杂网络逐渐被认识, 人们希望探索这些真实网络上的动力学过程^[1-4], 以便研究网络的动态演化. 复杂网络作为一种描述系统组元和组元之间关系的工具, 使研究者能够通过网络演化模型的探索来加强对真实网络认识的深度. 关于复杂网络拓扑结构和动力学特性的研究引起了人们广泛关注, 目前研究较多的动力学过程有传播、同步、博弈、相继故障、搜索等, 而搜索又是所有动力学的基本过程. 一个搜索过程简而言之就是粒子从初始节点出发, 沿着网络中的边到达目标节点. 但是, 大规模网络中两点之间的路径有很多条, 粒子要想在网络中快速到达目标节点, 就要找到节点对之间的较短乃至最短路径, 而这往往要依赖于网络的拓扑结构^[5]、节点的局部信息^[6,7]以及所使用的搜索算法^[8,9]等等. 优秀的搜索策略就是要花费很小的代价达到尽可能高的搜索效率. 文献 [10] 提出了一种单粒子和多粒子相结合的混合搜索模型, 既克服了单粒子搜索效率低下的缺点, 又降低了多粒子搜索的硬件代价;

Zhou^[11] 提出的混合导航模型通过对网络中加入智能路由器提高粒子的搜索效率. 这两种方法在无向网的搜索中表现出了优异的搜索性能. 但是对有向网, 由于路径具有指向性, 仿真结果显示有 50% 的可能性, 粒子会因“无路可走”而被迫终止搜索过程, 显然这是无法接受的. 因此设计经济有效的有向网络搜索策略是一个具有重要意义和挑战性的工作.

2011 年 Nature 杂志上发表了一篇关于复杂网络可控性 (network controllability, NC) 的文章, 作为一项开拓性工作, 迅速引起了人们的关注和研究^[12-15]. Liu 和 Barabasi^[12] 对线性时不变系统应用现代控制理论建立了复杂网络的可控性模型 (network controllability model, NCM), 提出了用结构可控性理论^[16] (Lin's structural controllability theory, LSCT) 和最小输入理论^[12] (minimum inputs theory, MIT) 求解最小驱动节点集的方法.

网络可控性是在某时刻能够独立控制有差异节点的状态变量, 最后使所有的节点处于期望的状态. 一个网络要能被控制, 意味着控制信号能被传输到每一个节点, NC 的本质就是控制信号的可达问题. NCM 启示我们可以依据网络可控性模型计算出的驱动节点 (driver nodes)、关键边 (critical

* 国家自然科学基金 (批准号: 60902094) 和全军军事学研究生课题 (批准号: 2010JY0072-046) 资助的课题.

[†] 通讯作者. E-mail: xiaoyandong08@gmail.com

links) 为粒子在搜索过程中快速找到一条可达路径, 同时又可以避免粒子“无路可走”的现象.

为了解决有向网的搜索效率和“无路可走”的问题, 本文提出了基于网络最大可控子图的导航搜索模型. 本文第 2 部分介绍了网络可控性的模型并给出了最大可控子图的定义; 第 3 部分提出了基于最大可控子图的导航搜索模型; 第 4 部分通过在随机网络和无标度网络上应用该模型, 检验算法的有效性; 第 5 部分总结全文.

2 网络的最大可控子图

2.1 网络可控性模型

网络控制技术的核心在于解决可行性和有效性两个问题: 可行性是研究一个网络到底能不能被控制; 有效性则回答了最少需要输入多少信号, 网络才能达到全局可控. NCM 中应用结构可控性定理回答了可行性问题, 而最小输入定理则解决了有效性问题.

对于线性时不变系统, 其动力学方程描述为

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad (1)$$

$\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_N(t))^T$ 表示 t 时刻网络中节点的状态, \mathbf{A} 为网络节点的邻接关系矩阵, \mathbf{B} 为控制信号与节点的连接关系矩阵, $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_M(t))^T$ 表示 t 时刻输入的控制信号的状态. 根据这些定义我们给出结构可控性判据.

定理 1 结构可控性定理 (Lin's structural controllability theorem)

下列条件是等价的:

- 1) 一个线性系统 $G(\mathbf{A}, \mathbf{B})$ 是结构可控的;
- 2) (i) $G(\mathbf{A}, \mathbf{B})$ 中不包含不可达 (inaccessibility) 节点,
(ii) $G(\mathbf{A}, \mathbf{B})$ 中不包含膨胀结构 (dilation);
- 3) $G(\mathbf{A}, \mathbf{B})$ 由 cacti 结构生成.

不可达节点是指外部输入的信号无法通过有向路径到达的节点. 如果 \mathbf{A} 的一个子集 S , 使得 $|T(S)| < |S|$, 则认为 \mathbf{A} 包含一个膨胀, $T(S)$ 是子集 S 的邻居集合. cacti 是既不含不可达节点又不含膨胀的网络结构, 若去除其中任意一条边将会导致不可达节点或膨胀的出现. 关于 cacti 的详细定义见文献 [12].

一般认为, 对于一个线性时不变系统, 如果它被证明满足结构可控性定理, 那么就认为它是一个

可控系统. NCM 的另一个贡献就是解决了网络可控性中最小驱动节点集的可计算性问题. 在 NCM 中已经被证明——根据最小输入定理, 网络的可控性问题可以转化成研究有向图的最大匹配问题 [12].

定理 2 最小输入定理 (minimum inputs theorem)

网络的最小输入集 (N_I) 等价于最小驱动节点集 (N_D). 如果网络是完美匹配, 最小输入集是网络中的任意一个节点; 否则, 它等于网络最大匹配后未被匹配的节点集. 定理 2 可用 (2) 式表示, $|M^*|$ 为有向网中最大匹配的节点数目:

$$|N_I| = |N_D| = \max\{1, N - |M^*|\}. \quad (2)$$

求解有向图的最大匹配, 首先要将有向图转化成二部图. 考虑一个有向图 $G(\mathbf{A})$, 与它对应的二部图为 $H(\mathbf{A}) = (V_A^+, V_A^-, \Gamma)$. V_A^+ 和 V_A^- 是二部图的两个点集, $V_A^+ = \{x_1^+, x_2^+, \dots, x_N^+\}$, $V_A^- = \{x_1^-, x_2^-, \dots, x_N^-\}$. Γ 是它的边集, $\Gamma = \{(x_i^+, x_j^-) | a_{ij} \neq 0\}$, 如果在 $G(\mathbf{A})$ 中节点 i 有边指向 j , 则在 $H(\mathbf{A})$ 中对应 x_i^+ 到 x_j^- 的有向边. 根据二部图的最大匹配算法, 计算出 $H(\mathbf{A})$ 中所有的最大匹配边. 最大匹配边也就是网络中控制信号到达节点的传输路径. 研究表明, 网络中对维持可控性发挥重要作用的最大匹配边有 $N - |N_D|$ 条.

定义 1 驱动节点 (driver nodes)

有最大匹配边指向的节点称为匹配节点, 没有最大匹配边指向的节点称为非匹配节点. 最大匹配中非匹配节点就是网络的驱动节点. 只要对驱动节点输入控制信号就能维持网络的全局可控性.

定义 2 控制节点 (control nodes)

对驱动节点输入控制信号的节点称为控制节点, 控制节点是系统之外的节点.

定义 3 关键边 (critical links)

如果一个边被移除, 我们需要增加驱动节点的数量去维持对网络的控制, 则这个边是关键边; 如果一个边的移除不影响当前的驱动节点集, 则这个边是冗余边; 如果一个边的移除虽然改变了网络中的驱动节点集, 但是没有增加驱动节点的数量, 则这个边是普通边. 显然网络中有许多不同的最大匹配边集合, 这些集合是由全部的关键边和部分冗余边组成的, 每一个集合中元素的个数为 $N - |N_D|$. 本文中只需找到网络中的一个最大匹配边集合, 并将这个最大匹配边集合狭义地定义为关键边, 其余边

称为非关键边. 后文中关键边如无特指就是狭义的关键边.

2.2 最大可控子图

根据网络可控性模型, 对于任意拓扑结构的网络都可以计算出其最小的驱动节点集、关键边、冗余边和普通边. 图 1(a) 是一个包含 14 个节点的原始网络 $G(\mathbf{A})$; 图 1(b) 为网络中的两个 cacti 结构, 虚线为网络中的关键边; 图 1(c) 为网络中驱动节点的分布, 方框为驱动节点, 外部输入为控制节点, 灰色边为网络中的非关键边. 由 LSCT 可知, 关键边

用来把节点连接起来组成链状或环状结构. 通过这些环状和链状路径, 控制信号就能被传输到每一个节点上.

定义 4 最大可控子图 (subnet of maximum controllability, SMC)

一个线性系统 $G(\mathbf{A}, \mathbf{B})$ 中存在一个子图 S , 在 S 中至多只需对一个驱动节点输入控制信号就可以维持 S 中所有节点可控, 且去除 S 中的任意一条边都会增加驱动节点, 满足这样条件的子图称为最大可控子图. 图 1(c) 中的阴影区域就是网络中的 4 个最大可控子图.

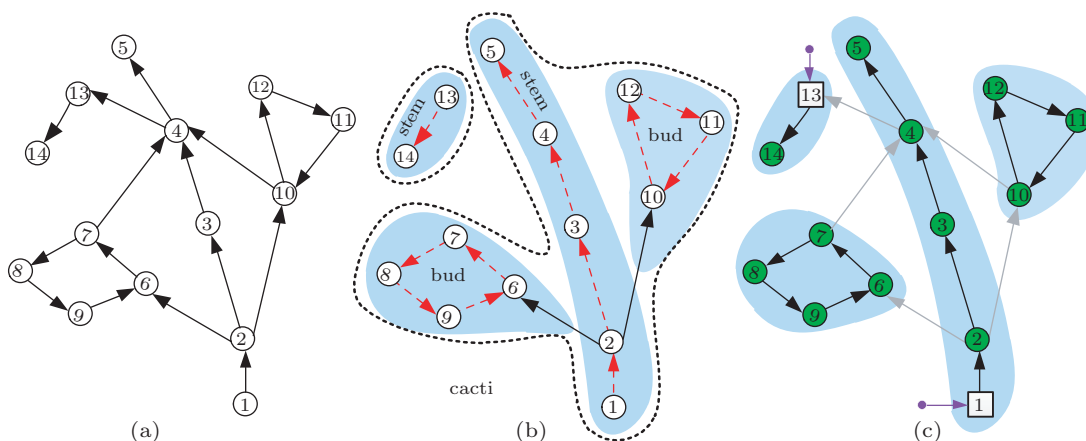


图 1 网络中的最大可控子图 (a) 有向网络 $G(\mathbf{A})$; (b) $G(\mathbf{A})$ 中包含的两个 cacti 结构; (c) $G(\mathbf{A})$ 中驱动节点分布和 4 个最大可控子图

从定义可知, 最大可控子图分为链状和环状可控子图两种, 其中的边都是关键边. 对于环状可控子图, 从任意节点出发都可以到达其他的所有节点; 对于链状的最大可控子图, 只能从驱动节点出发才可以到达其他的节点. 如果将网络中所有的 SMC 首尾连接起来能解决粒子因“无路可走”而结束搜索的问题, 将极大提高粒子搜索的成功率. Wang 等^[17] 分析了将所有链状 SMC 首尾连接起来是一种高效的优化网络可控性的方法.

3 基于最大可控子图的导航模型

3.1 随机游走

网络上的节点在不附加任何条件的情况下, 节点的搜索实质上是一个布朗粒子随机游走 (random walk, RW) 的问题. 从概率论的角度讲, 随机游走是一个一阶马尔可夫过程, 一个粒子从一个节点跳到另一个节点就发生了一次概率转移. 假设 $t = 0$ 时

刻, 一个粒子从节点 a 出发随机游走之后到达搜索目标节点 b , 这个过程的概率转移方程就表达为

$$\begin{aligned}
 p_{ab}(t) &= p_{aj_1}(1) \cdot p_{j_1 j_2}(2) \cdots p_{j_s b}(t) \\
 &= \sum_{j_1 j_2 \cdots j_s} \frac{A_{aj_1}}{k_a} \cdot \frac{A_{j_1 j_2}}{k_{j_1}} \cdots \frac{A_{j_s b}}{k_{j_s}}, \quad (3)
 \end{aligned}$$

其中 $p_{ab}(t)$ 为粒子在时刻 t 到达节点 b 的概率, k_i 表示节点 i 的度. \mathbf{A} 为 $N \times N$ 的邻接矩阵, A_{ij} 表示节点 i 和 j , 若之间有边相连, 则 $A_{ij} = 1$; 否则, $A_{ij} = 0$. 若用矩阵 $\mathbf{P}(P_{ij} = A_{ij}/k_i, k_i = \sum_{j=1}^N A_{ij})$ 表示粒子在网络中随机游走的概率转移矩阵, $\pi_s^x(t)$ 表示粒子 s 在 t 时刻处于节点 x 的概率. 那么 $\pi_s^x(0)$ 是 $N \times 1$ 的向量, 表示初始时刻粒子 s 处在节点 x 处, 则第 x 位为 1, 其余全为 0. 当时刻 t 时, 我们可以得到粒子的位置概率向量为

$$\pi_s(t) = \mathbf{P}^T \pi_s(t-1). \quad (4)$$

根据 $\pi_s^x(t)$ 可以确定粒子下一跳的位置. 文献 [18] 计算了粒子随机游走的首达时间期望.

随机搜索的缺点显而易见. 粒子在网络中游走就像布朗运动一样, 搜索效率很低. 尤其在有向网络中, 由于只有单一指向的路径, 粒子在搜索过程中很容易发生因“无路可走”而结束搜索的问题. 如图 1(a) 的 $G(\mathbf{A})$ 网络中, 当粒子跳转到节点 5 时, 由于没有路径可以跳转, 粒子就只能结束搜索. 假设在图 1(a) 所示的网络拓扑中, 粒子初始位于节点 6

处, 要搜索到目标节点 14. 仿真得到了 $G(\mathbf{A})$ 网络中游走的首达时间概率分布如图 2 所示, 其中 $E(T)$ 为粒子游走的首达时间期望. 从图 2 得知, 在原始的 $G(\mathbf{A})$ 网络中, 10^6 次实验中有 50.086% 是因为粒子由于没有路径再跳转而无法到达目标节点, 实验中认为这种情况的首达时间是 $+\infty$, 显然这样的搜索结果对有向网而言是不能接受的.

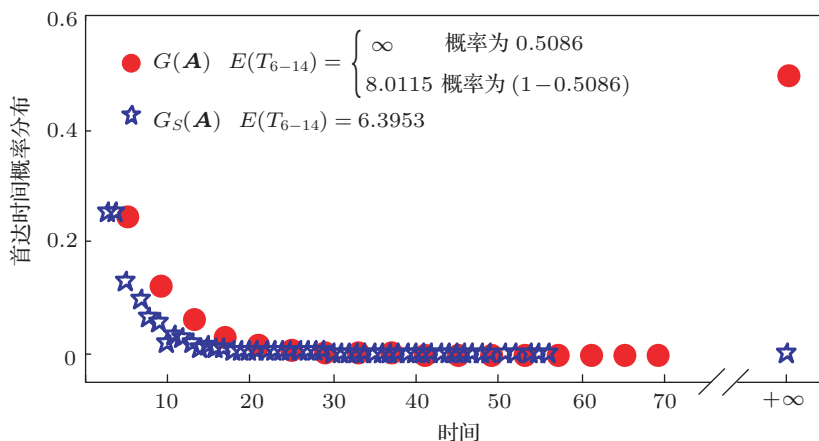


图 2 $G(\mathbf{A})$ 和 $G_S(\mathbf{A})$ 网络中首达时间概率分布, $E(T)$ 为粒子游走的首达时间期望, 实验仿真 10^6 次

3.2 模型理论

搜索问题作为网络动力学的基础, 从最初的穷举方法^[19]到局部搜索策略^[20]以及具有智能作用的自适应路由机制^[11], 虽然很好地提高了无向网的搜索效率, 但是对于有向网一直没能很好解决“无路可走”和搜索效率的问题. Liu 和 Barabasi^[12]的可控性模型中控制信号从驱动节点输入, 信号沿着网络中的关键边到达网络中的每一个节点. 根据这样的思想, 如果将控制信号看作是粒子, 网络中的关键边就为粒子游走提供了一条可达路径. 从第二章提出的可控子图的概念看到, 只要将网络中所有的可控子图首尾连接起来就构成了一个大的环状可控子图, 那么网络中就是一个全联通的网络, 任意两个节点之间都可以互相到达, 将这种基于最大可控子图的有向搜索网络记为 $G_S(\mathbf{A})$.

图 3 表示了生成 $G_S(\mathbf{A})$ 的两种具体过程. 假设为图 1(a) 中的有向网络 $G(\mathbf{A})$, 根据网络可控性模型计算出的 4 个 SMC 如图 1(c) 所示. 将网络中的 SMC 的个数记为 N_M , 其中环状 SMC 的数量为 N_C , 驱动节点的数量记为 N_D , 向网络中添加的边的数量为 N_A . 图 3(a) 是先将 $G(\mathbf{A})$ 中所有的链状 SMC 首尾相连成一个大的环状 SMC (单向虚线表示添加的边), 再将 $N_C - 1$ 个环状 SMC 连接到剩余的一

个环状 SMC 上 (双向虚线表示添加的边). 此时 N_A 满足

$$N_D \leq N_A \leq N_D + 2(N_C - 1), \quad (5)$$

如果两个环状 SMC 之间本来就存在边, 则不需要额外添加边了. 图 3(b) 是另一种加边方法, 将 $G(\mathbf{A})$ 中所有的最大可控子图用单向边首尾连接起来, 那么只需向网络中添加 N_A 条边,

$$N_A = N_M = N_D + N_C, \quad (6)$$

这虽然是最少的加边方法, 但这将带来搜索路径的增长, 因为粒子在两个链状 SMC 之间跳转有可能要经过多个环状 SMC 传递. 如在图 3(b) 中, 如果粒子初始位于节点 13 处, 要跳转到节点 5 处, 要经过两个环状 SMC. 但是图 3(a) 的加边方法有效避免了两个链状 SMC 之间跳转要经过其他环状 SMC 传递的问题, 而且图 3(a) 中至多比图 3(b) 多添加 $N_C - 2$ 条边. 显然图 3(a) 的基于最大可控子图的有向搜索网络 $G_S(\mathbf{A})$ 是一种经济有效的加边方式, 因此 $G_S(\mathbf{A})$ 采用图 3(a) 加边方式.

图 2 显示了在初始 $G(\mathbf{A})$ 网络和新生成的 $G_S(\mathbf{A})$ 网络上进行实验的仿真结果, 仿真结果说明 $G_S(\mathbf{A})$ 的搜索网络解决了粒子在有向网中因“无路可走”而结束搜索的情况. 对于 $G_S(\mathbf{A})$ 的搜

索网络, 粒子搜索的成功率大大提高, 同时粒子的平均首达时间减小了. 但是对比图 2 发现, 在能够成功搜索到目标节点的情况下, 两种搜索网络的首达时间的概率分布的趋势并没有发生变化, 概率分布的峰值所对应的首达时间也没有明显减少. 这说明在 $G_S(\mathbf{A})$ 的搜索网络中仅提高了粒子的搜索成

功率, 而没能给粒子提供一条快速到达目标的路径, 粒子在网络中跳转依然目标性不强, 搜索效率并没有达到最优. 因此, 对一个有向网络的搜索问题, 不仅要解决粒子游走过程中“无路可走”的问题, 又要避免粒子在游走过程中随机性, 快速找到最便捷的路径.

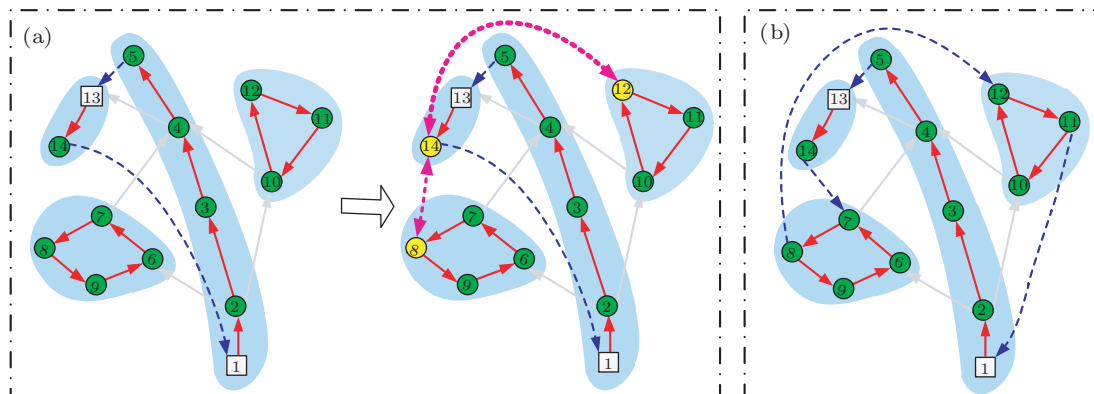


图 3 两种基于最大可控子图的搜索模型的加边示意图 (a) 将网络中链状 SMC 连接成环状 SMC, 再把 $N_C - 1$ 个环状 SMC 连接到剩余的一个环状 SMC 上生成有向搜索网络 $G_S(\mathbf{A})$; (b) 将所有的 SMC 用单向边首尾连接起来生成有向搜索网络 $G_S(\mathbf{A})$

Zhou^[11] 提出了一种混合导航机制 (mixing navigation mechanism, MNM) 很好地提高了无向网中粒子的搜索效率. MNM 通过向网络中添加一定数量的智能路由器来优化网络的搜索效率. MNM 的主要思想是: 当粒子经过路由器时, 它会沿目标节点的最短路径方向到达下一个节点, 从而降低了搜索过程的随机性. 其中智能路由器的数量和位置也对模型的平均首达时间有很大影响. 受到 MNM 模型的启发, 本文对前面提出的有向搜索网络 $G_S(\mathbf{A})$ 进行了修改, 在网络中加入“智能路由器”, 使基于最大可控子图的搜索模型成为导航模型, 为粒子提供快速到达的路径. 基于最大可控子图的导航模型 (navigation model based subnet of maximum controllability, NMSMC) 算法的步骤如下. 粒子不能在非关键边上游走, 相当于 $G_N(\mathbf{A})$ 的导航网络中删除非关键边, 仅保留关键边. 导航边连接的节点称为导航点, 新加入的导航边上存储有所连接的两个环状最大可控子图的节点集合. 若粒子在导航点, 根据导航边提供的信息, 判断目标粒子是否在该 SMC 上, 若在则继续在该 SMC 中搜索, 若不在则跳转到另一个 SMC 中继续判断, 直到根据导航边提供的信息找到目标节点的位置.

基于 NMSMC 算法描述如下:

输入: 有向网络 $G(\mathbf{A})$

输出: 基于 NMSMC 的导航搜索网络 $G_N(\mathbf{A})$,

粒子的首达时间 T_{ab}

过程:

- (i) 计算最大可控子图 应用网络可控性计算出有向网络 $G(\mathbf{A})$ 中所有的最大可控子图集合;
- (ii) 标记非关键边 将网络中的非关键边进行标记, 粒子将不能再在被标记过的边上跳转;
- (iii) 生成环状 SMC 将网络中所有链状的 SMC 首尾连接起来构成一个较大的环状 SMC;
- (iv) 生成 NMSMC 的有向导航网络 根据 $G_S(\mathbf{A})$ 中的加边方法, 用双向导航边将环状 SMC 连接起来, 生成 NMSMC 的有向导航网络 $G_N(\mathbf{A})$;
- (v) 判断路径 粒子沿着网络中的关键边跳转, 如果在跳转过程中发现目标节点, 则结束搜索; 否则, 在导航点处根据导航边的路由表判断下一步要跳转的 SMC;
- (vi) 计算粒子的搜索时间 T_{ab} .

4 模型应用

上面的实验只关注指定节点之间的平均首达时间, 本节的实验将选择网络中任意两个节点作为出发和目标节点, 通过在 ER 和 SF 网络上进行实验, 检验 NMSMC 算法的有效性. 在节点数 $N = 300$ 的网络中, 每次随机抽取一对节点作为起始节点和目标节点, 共运行 10^6 次后取平均值, 得出全网的

平均搜索时间. 用 $T(G_N(\mathbf{A}))$ 和 $L(G_N(\mathbf{A}))$ 分别表示 $G_N(\mathbf{A})$ 的全网搜索时间和网络的平均最短路径.

图 4 所示为在 ER 网络上的实验结果. 首先比较了加入导航点前后带来的搜索效率的变化. 从图 4(c) 中发现, 在 $G_N(\mathbf{A})$ 和 $G_S(\mathbf{A})$ 的搜索模型中, $T(G_N(\mathbf{A}))$ 和 $T(G_S(\mathbf{A}))$ 在 $\langle k \rangle = 1$ 时相同, 但随着 $\langle k \rangle$ 增加, $T(G_N(\mathbf{A}))$ 和 $T(G_S(\mathbf{A}))$ 出现了两种截然不同的变化趋势, $T(G_S(\mathbf{A}))$ 越来越大, 而 $T(G_N(\mathbf{A}))$ 越来越小, 向 $L(G_N(\mathbf{A}))$ 靠近, 说明基于 NMSMC 的导航模型表现出了非常优秀的搜索性能. 从图 4(a) 和 (b) 中发现, 随着 $\langle k \rangle$ 增加, 驱动节点的比例在下降, 网络的可控性越来越好, 此时网络中的环状 SMC 增多, 使得网络中导航点也随之增加. 导航点越多, 全网的搜索时间就越小. 图 4(b) 则说明达到这样的搜索效果, 付出的代价并不是很高.

随着 $\langle k \rangle$ 的增加, 需要向网络中添加的边的数量在不断减小, 尤其是在 $\langle k \rangle$ 较高时, 需要向网络中添加的边的数量很少. 此外, 由于网络中导航点的数量与搜索时间是成反比的, 实验结果却显示, 以不到节点总数 2% 的导航点却达到了接近平均最短路径的搜索效率. 从费效比上说明在最大可控子图上加边和部署导航点是一种极为有效的搜索策略.

基于 NMSMC 生成的导航网络 $G_N(\mathbf{A})$ 中, 如果假设粒子可以在网络的非关键边上跳转, 把这样的网络记为 $G'_N(\mathbf{A})$. $G'_N(\mathbf{A})$ 中全网的平均搜索时间和网络的平均最短路径分别记为 $T(G'_N(\mathbf{A}))$ 和 $L(G'_N(\mathbf{A}))$, 它满足

$$\begin{aligned} 1 &\leq L(G'_N(\mathbf{A})) \leq T(G'_N(\mathbf{A})) \\ &\leq L(G_N(\mathbf{A})) \leq T(G_N(\mathbf{A})), \end{aligned} \quad (7)$$

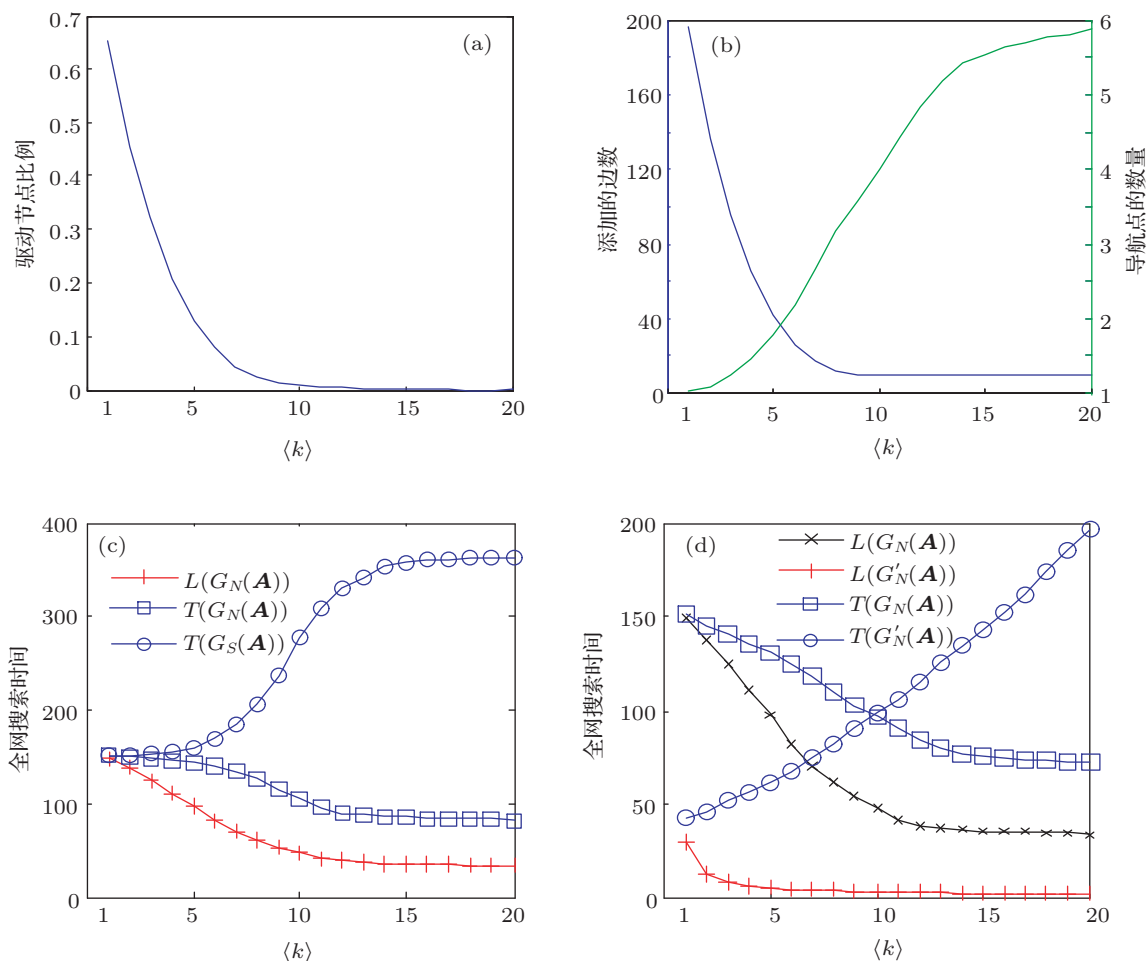


图 4 ER 网络上搜索的实验结果, $N = 300$ (a) 随 $\langle k \rangle$ 增加, 网络中驱动节点的比例变化; (b) 在 $G_N(\mathbf{A})$ 导航网络中, 随 $\langle k \rangle$ 增加, 网络中要添加边的数量和部署导航点的数量; (c) 随 $\langle k \rangle$ 增加, 在 $G_N(\mathbf{A})$ 和 $G_S(\mathbf{A})$ 中, 全网的平均搜索时间变化, $L(G_N(\mathbf{A}))$ 为 $G_N(\mathbf{A})$ 的平均最短路径; (d) 随 $\langle k \rangle$ 增加, 在 $G_N(\mathbf{A})$ 和 $G'_N(\mathbf{A})$ 中, 全网的平均搜索时间变化, $L(G'_N(\mathbf{A}))$ 为 $G'_N(\mathbf{A})$ 的平均最短路径, $L(G_N(\mathbf{A}))$ 为 $G_N(\mathbf{A})$ 的平均最短路径

平均最短路径随着向网络中添加边而不断变小,直到网络中任意两个节点之间都有边相连,网络的平均最短路径最小为 1. $G_N(\mathbf{A})$ 是由许多小环组成的一个大环网络,且 $G_N(\mathbf{A})$ 网络中粒子可以跳转的边要远小于 $G'_N(\mathbf{A})$,因此 $G_N(\mathbf{A})$ 的平均最短路径要大于 $G'_N(\mathbf{A})$. 如果粒子在 $G'_N(\mathbf{A})$ 中非导航点上能尽可能地找到节点间的最短路径, $T(G'_N(\mathbf{A}))$ 将无限接近于 $L(G'_N(\mathbf{A}))$,而远小于 $L(G_N(\mathbf{A}))$ 和 $T(G_N(\mathbf{A}))$. (7) 式给出了 $G_N(\mathbf{A})$ 和 $G'_N(\mathbf{A})$ 的全网搜索时间和平均最短路径的关系. 图 4(d) 的实验结果验证了这种关系,发现 (7) 式只适用 $\langle k \rangle$ 较小的网络中, $T(G'_N(\mathbf{A}))$ 随着 $\langle k \rangle$ 的增大不断增加. 当 $\langle k \rangle \leq 10$ 时, $L(G'_N(\mathbf{A})) < T(G'_N(\mathbf{A})) < T(G_N(\mathbf{A}))$; 当 $\langle k \rangle \geq 11$ 时, $T(G'_N(\mathbf{A})) > T(G_N(\mathbf{A})) > L(G_N(\mathbf{A})) > L(G'_N(\mathbf{A}))$. 这是由于 $\langle k \rangle$ 较小时,边的数量较少,粒

子在非导航节点处随机跳转的选择较少,同时粒子在 $G'_N(\mathbf{A})$ 中有比 $G_N(\mathbf{A})$ 中更多的“捷径”可以跳转,使得 $T(G'_N(\mathbf{A}))$ 小于 $T(G_N(\mathbf{A}))$; 但是当 $\langle k \rangle$ 较大时,网络变得稠密,边的数量很多,同时 NMSMC 模型中导航点的数量并没有随 $\langle k \rangle$ 增加很多,此时粒子在 $G'_N(\mathbf{A})$ 中非导航点上会沿着其他路径随机跳转,增加了搜索时间. 图 4(d) 的实验结果说明了 $G_N(\mathbf{A})$ 的导航搜索策略要优于 $G'_N(\mathbf{A})$,尤其是在稠密的网络中.

图 5 所示在 $\gamma = 2.5, 3, 4$ 的 SF 网络上的实验结果. SF 网络的 $T(G_N(\mathbf{A}))$, $T(G_S(\mathbf{A}))$ 和 $T(G'_N(\mathbf{A}))$ 的变化趋势与 ER 网络一致,说明仅在网络关键边上有导航的游走是一种极为有效的搜索策略,避免了粒子在网络中游走的随机性,尤其适合解决平均度很高的网络的搜索问题. 实验结果还显示,相同

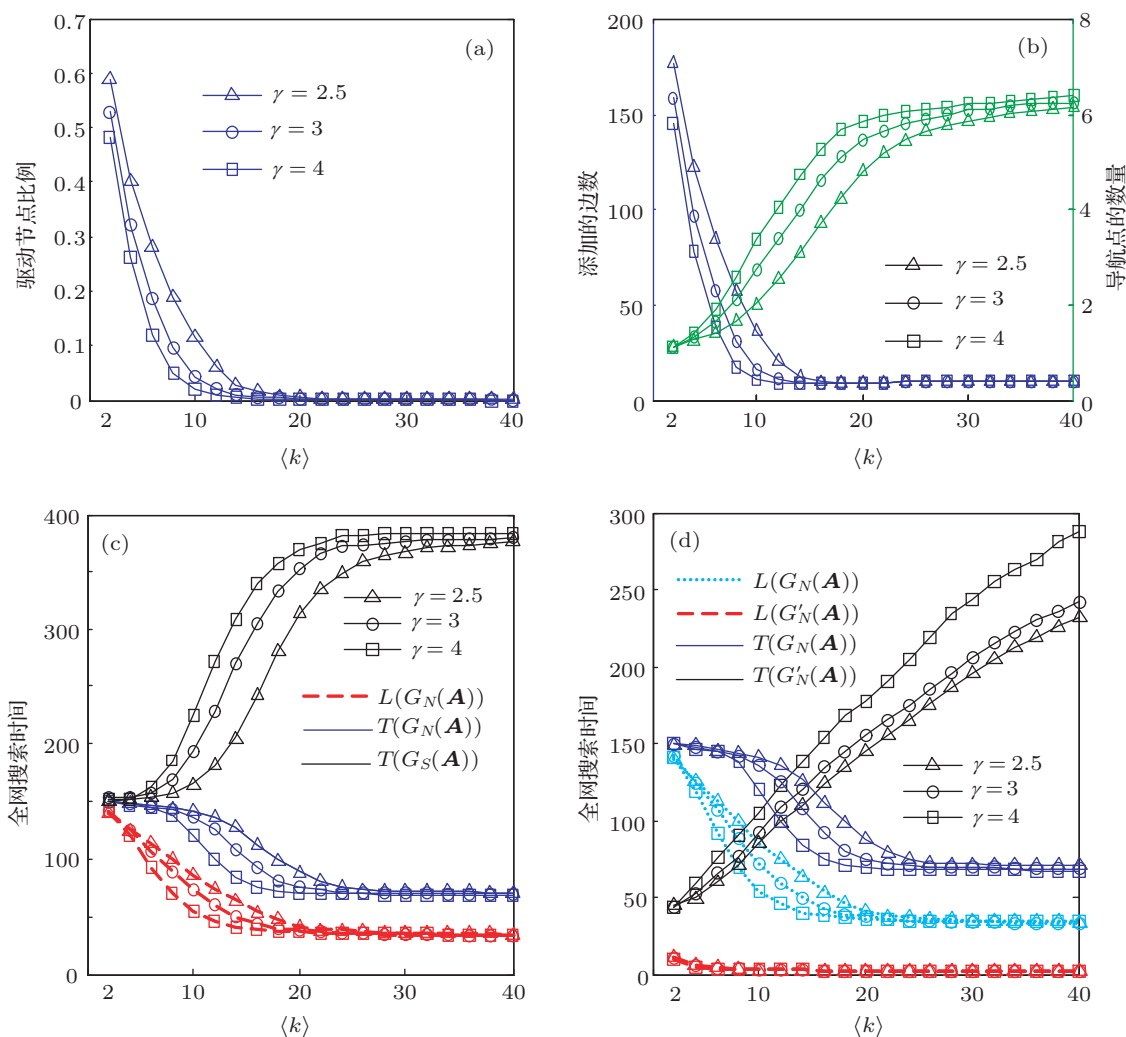


图 5 SF 网络上搜索的实验结果, $N = 300$, $\gamma = 2.5, 3, 4$ (a) 随 $\langle k \rangle$ 增加,网络中驱动节点的比例变化; (b) 在 $G_N(\mathbf{A})$ 导航网络中,随 $\langle k \rangle$ 增加,网络中要添加的边的数量和导航点的数量; (c) 随 $\langle k \rangle$ 增加,在 $G_N(\mathbf{A})$ 和 $G_S(\mathbf{A})$ 中,全网的平均搜索时间变化, $L(G_N(\mathbf{A}))$ 为 $G_N(\mathbf{A})$ 的平均最短路径; (d) 随 $\langle k \rangle$ 增加,在 $G_N(\mathbf{A})$ 和 $G'_N(\mathbf{A})$ 中,全网的平均搜索时间变化, $L(G'_N(\mathbf{A}))$ 为 $G'_N(\mathbf{A})$ 的平均最短路径, $L(G_N(\mathbf{A}))$ 为 $G_N(\mathbf{A})$ 的平均最短路径

平均度下, γ 越大, $T(G_N(\mathbf{A}))$ 越小, 而 $T(G_S(\mathbf{A}))$ 和 $T(G'_N(\mathbf{A}))$ 越大, 这是由于网络可控性不同造成的. 从图 5(a) 和 (b) 中发现, SF 网络在相同 $\langle k \rangle$ 时, 随着 γ 增大, 网络可控性越好, 使得网络中分布的导航点数量增加, 减少了粒子在网络中的搜索时间. 对比 ER 和 SF 的实验结果也可以发现, 由于 SF 网络的可控性要差于 ER 网络, 在相同平均度下使得 SF 网络中的全网搜索时间要大于 ER 网络.

NMSMC 模型在 ER 和 SF 有向网络中表现出优异的搜索性能. 从费效比上看, 仅用节点总数 2% 左右的导航点和添加尽量少的边, 就使全网搜索时间接近导航网络的平均最短路径, 说明基于最大可控子图加边和部署导航点策略非常有效. 此外, 全网搜索时间与网络的可控性的优劣紧密相关, 可

控性越好, 使网络中导航节点分布越多, 越能提高网络的搜索效率.

5 结论

在 Liu 和 Barabasi 提出的网络可控性的框架基础上, 本文提出了最大可控子图的概念. 为了解决有向网络搜索问题中存在的搜索效率低下和“无路可走”的问题, 本文提出了一种基于最大可控子图的导航搜索模型. 通过分析和对比发现, NMSMC 是一种极为有效的搜索策略, 不仅能通过添加尽可能少的边来解决粒子“无路可走”的问题, 而且在最大可控子图上加边和部署导航点, 花费极小的代价就能达到接近导航网络平均最短路径的搜索效率.

-
- [1] Barabási A L, Albert R 2002 *Rev. Mod. Phys.* **74** 47
- [2] Newman M E J 2003 *SIAM Rev.* **45** 167
- [3] Li J, Wang B H, Jiang P Q, Zhou T, Wang W X 2006 *Acta Phys. Sin.* **55** 4051 (in Chinese) [李季, 汪秉宏, 蒋品群, 周涛, 王文旭 2006 物理学报 **55** 4051]
- [4] Feng C F, Xu X J, Wu Z X, Wang Y H 2008 *Chin. Phys. B* **17** 1951
- [5] Adamic L A, Lukose R M, Puniyani A R, Huberman B A 2001 *Phys. Rev. E* **64** 046135
- [6] Kim B J, Yoon C N, Han S K, Jeong H 2002 *Phys. Rev. E* **65** 027103
- [7] Thadakamalla H P, Albert R, Kumara S R T 2005 *Phys. Rev. E* **72** 066128
- [8] Song Q, Wang X F 2011 *IEEE Trans. Intellig. Transport. Syst.* **12** 132
- [9] Bauer R, Dellling D, Sanders P, Schieferdecker D, Schultes D, Wagner D 2010 *ACM J. Experiment. Algorithms* **15** 1
- [10] Pu C L, Pei W J, Wang S P 2010 *Acta Phys. Sin.* **59** 103 (in Chinese) [濮存来, 裴文江, 王少平 2010 物理学报 **59** 103]
- [11] Zhou T 2008 *Physica A* **387** 3025
- [12] Liu Y Y, Slotine J J, Barabási A L 2011 *Nature* **473** 167
- [13] Müller F J, Schuppert A 2011 *Nature* **478** E4
- [14] Egerstedt M 2011 *Nature* **473** 158
- [15] Lü T Y, Piao X F, Xie W Y, Huang S B 2012 *Acta Phys. Sin.* **61** 170512 (in Chinese) [吕天阳, 朴秀峰, 谢文艳, 黄少滨 2012 物理学报 **61** 170512]
- [16] Lin C T 1974 *IEEE Trans. Automat. Control* **19** 201
- [17] Wang W X, Ni X, Lai Y C, Grebogi C 2012 *Phys. Rev. E* **85** 026115
- [18] Zhou H J 2003 *Phys. Rev. E* **67** 041908
- [19] Dijkstra E W 1959 *Numer. Math.* **1** 269
- [20] Wang W X, Yin C Y, Yan G, Wang B H 2006 *Phys. Rev. E* **74** 016101

A navigation search model based on subnet of maximum controllability*

Xiao Yan-Dong[†] Lao Song-Yang Hou Lü-Lin Bai Liang

(College of Information Systems and Management, National University of Defense Technology, Changsha 410073, China)

(Received 6 May 2013; revised manuscript received 24 September 2013)

Abstract

In this paper, we propose a concept of subnet of maximum controllability based on the model of network controllability, and set up the navigation search model based on the subnet of maximum controllability, called NMSMC. The strategy of adding links that is based on the subnet of maximum controllability is to solve, with the minimum cost, the terminating search, the problem that arises from no way for particles to search in the directed network. Based on the subnet of maximum controllability to deploy navigation nodes, the search time of the whole network can be made close to the average shortest path of the navigation network, which the number of navigation nodes is only 2% of the total nodes. The experimental results of the ER and SF networks show that the search efficiency is strongly correlated with the network controllability. The better the controllability, the less the adding links are, which can lead to the fact that the more the navigation nodes are distributed in the network, the more the search efficiency of the network can be enhanced.

Keywords: navigation search, directed network, network controllability

PACS: 89.75.Hc, 89.75.Fb, 02.30.Yy

DOI: 10.7498/aps.62.248901

* Project supported by the National Natural Science Foundation of China (Grant No. 60902094) and the Military Science Foundation of China (Grant No. 2010JY0072-046).

[†] Corresponding author. E-mail: xiaoyandong08@gmail.com