

空间网络上的随机游走*

抖斐玲 胡延庆 黎勇 樊瑛 狄增如†

(北京师范大学管理学院系统科学系, 北京师范大学复杂性研究中心, 北京 100875)

(2011年11月21日收到; 2012年2月20日收到修改稿)

本文以一维均匀环为基础, 通过添加有限数量的长程连接构造出了一维有限能量约束下的空间网络, 环上任意节点 i 与 j 之间存在一条长程连接的概率满足 $p_{ij} \propto d_{ij}^{-\alpha}$ ($\alpha \geq 0$), 其中 d_{ij} 为节点 i 与 j 之间的网格距离, 并且所有长程连接长度总和受到总能量 $\Lambda = cN$ ($c \geq 0$) 的约束, N 为网络节点总数. 通过研究该空间网络上的随机游走过程, 存在最优幂指数 α_0 使得陷阱问题的平均首达时间最短. 进一步研究发现, 平均首达时间与网络规模 N 之间存在着幂律关系, 随着网络规模 N 和总能量 Λ 的增加, 最优幂指数 α_0 单调增加, 并趋近最优值 1.5.

关键词: 空间网络, 随机游走, 能量约束, 平均首达时间

PACS: 89.75.Hc, 89.75.Fb

1 引言

近年来, 复杂网络已经在数学、物理、生物、管理、计算机等诸多领域内引起了一股研究热潮. 除了网络的小世界特性、无标度特性等倍受关注的网络结构性质之外, 网络上的动力学行为也是一个研究热点^[1-3]. 其中, 随机游走就是诸多动力学过程中最基本, 同时也是研究历史最为悠久的一种, 一直以来为众多学者广泛研究^[4-9].

复杂网络上的随机游走是指以网络节点为载体, 按照一定概率从网络上任一节点转移到与之有连接的其他节点的状态转移过程. 大量研究结果表明, 网络上的随机游走不但与中心性 (random walk centrality)^[4]、网络平均距离^[10]、生成树数目^[11]等网络自身结构性质有关, 而且与网络上运输^[12]、导航^[13-15]、疾病传播^[16]、策略选择^[17]等其他动力学过程也存在密切关系.

目前, 作为研究网络上随机游走的载体网络主要涵盖了随机网、规则网、无标度网、分形或伪分形网、树形结构、以及一些人工构造的自相似

网络^[18-20]. 值得注意的是, 这些网络都主要着眼于网络的拓扑结构. 虽然这样的简化能够对很多复杂系统进行有效分析, 然而当元素之间的地理距离成为不可忽略的因素时, 网络的空间结构就可能会对整个系统起到非常重要的影响. 举例来讲, 自然界中广泛存在的人际关系网络, 血管网络, 神经网络, 电力网络, Internet 网络, 航空网络, 铁路网络等就因为受到空间地理位置因素的影响而具有明显的空间结构. 在这类网络上, 每个节点都有自身固有的空间位置, 而且其空间结构往往对网络的特征和功能有不可忽视的作用.

目前, 复杂网络上的随机游走研究内容主要包括: 网络的结构特性对随机游走的影响; 利用网络上随机游走探索网络结构特性, 并解决诸如划分社团结构、优化目标搜索、探索未知路径等问题; 也包括讨论网络上随机游走过程的首达或到达时间、首返或返回时间等问题: 即从给定网络节点 i 首达或到达节点 j 所需要的时间, 或者从网络节点 i 首次返回节点 i 需要的时间. 而在诸多有关随机游走的研究问题中, 陷阱问题颇具代表性且受到广泛关注^[20]. 所谓陷阱问题就是将网络中的某

* 中央高校基本科研业务费专项资金、国家自然科学基金 (批准号: 60974084, 61174150) 和 NCET-09-0228 资助的课题.

† E-mail: zdi@bnu.edu.cn

个节点作为吸收点, 考虑从其他节点出发以随机游走的方式到达这个陷阱所需的平均时间, 并且在此基础上讨论网络结构性质与平均首达时间 (MFPT) 的关系 [21-24].

本文将以空间网络为载体, 讨论引入空间地理位置因素后的网络上的随机游走过程. 进一步, 我们还将讨论空间结构性质对网络上的随机游走过程的影响. 为了简化, 我们以一维均匀环作为构建空间网络的基础. 假定该一维均匀环上的任意节点 i 都与其左右两个网格距离为 1 的邻居节点之间存在短程连接. 在此基础上, 我们以概率 $p_{ij} \propto d_{ij}^{-\alpha}$ ($\alpha > 0$) 在节点 i 与 j 之间添加长程连接, 其中 d_{ij} 指节点 i 和 j 之间的一维网格距离. 不难看出, 当节点 i 和 j 之间的网格距离越远, 那么它们之间存在一条长程连接的概率将越小, 即环上的节点之间存在长程连接的概率受空间地理位置因素的影响. 在实际情况中, 长程连接的增加通常意味着能量的消耗. 以改建地铁网络为例, 添加一条新的地铁线路将意味着巨大的成本, 而通常只有有限的资金预算. 为了更切合实际, 本文也引入了总能量约束, 即在该一维环上添加的长程连接长度总和满足 $\Lambda = c \cdot N$ 的限制, 其中 N 为节点总数, $c > 0$ 为常数. 通过讨论该一维有限能量约束下的空间网络上的随机游走过程, 我们发现存在一个最优的空间结构性质参数 α_0 使得该空间网络上的陷阱问题平均首达时间最短. 进一步, 我们研究了网络规模 N 以及总能量约束 Λ 对该最优幂指数 α_0 的影响.

2 空间网络模型与最优幂指数

2.1 建立一维能量约束下的空间网络

我们从节点总数为 N 的一维均匀环出发, 将各节点按环的连接顺序顺次编号 $1, 2, \dots, N$, 该一维环上的任意节点 i 与左右两个相邻节点之间都存在网格距离为 1 的短程连接. 另外, 我们以概率 $p_{ij} \propto d_{ij}^{-\alpha}$ 的幂律方式在节点 i 与节点 j 之间添加长程连接, 其中 d_{ij} 为节点 i 和 j 的一维网格距离, α 是幂指数, 文中取 $0 \leq \alpha \leq 3$. 给定总能量约束 Λ , 我们将以粗粒化的方法建立一维有能量约束下的一维空间网络 [25]. 具体步骤如下:

1) 将一维环的所有 N 个节点分别与其最近的

两个邻居节点建立短程连接, 并选择粗粒化网络的参数 W 值;

2) 随机等概率选择一个节点 i , 以概率 $P(r) = \beta \cdot r^{-\alpha}$ 产生距离 r ($2 \leq r \leq N_{\max}$), 其中 β 为归一化因子, 满足 $\sum_{r=2}^{N_{\max}} P(r) = 1$;

3) 对于节点 i , 剩余节点将按照粗粒化参数 W 分成 $\frac{N_{\max}}{W}$ 个部分 (N_{\max} 是粗粒化前最大网格距离), 其中第 m 部分 ($1 \leq m \leq \frac{N_{\max}}{W}$) 包含粗粒化前与该节点网格距离为 $(m-1)W+1, (m-1)W+2, \dots, mW$ 的所有节点;

4) 若距离 r 落入 i 节点粗粒化后的第 m 部分, 则等概率地从 i 节点的第 m 部分中随机选择一个节点 j , 若 i 和 j 之间没有连接, 则在这两个节点之间建立一条长程连接, 考虑到能量消耗正比于节点的网格距离, 我们假定该长程连接使总能量消耗 r ;

5) 重复步骤 2) 至 4), 直到总能量 Λ 耗尽.

图 1 给出了一维有限能量约束下的空间网络的示意图. 由于一维均匀环节点的对称性, 环上所有节点是同质的, 当添加长程连接后, 在概率意义下, 环上各节点仍然同质.

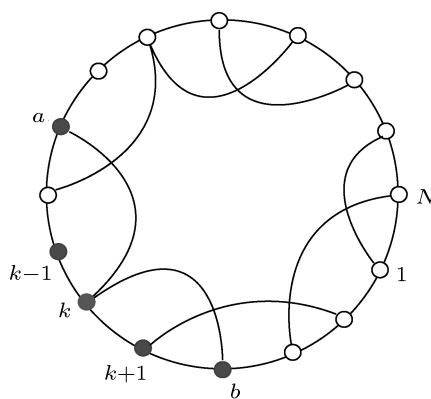


图 1 网络示意图

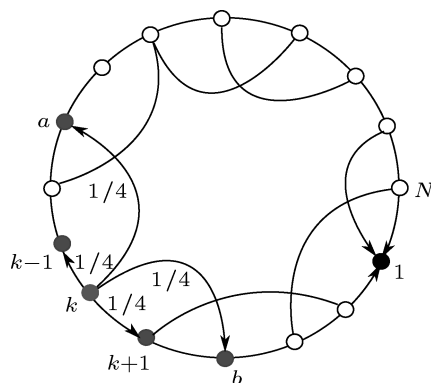


图 2 无规游走、陷阱点示意图

2.2 蒙特卡洛模拟

考虑在通过前述步骤构建的空间网络上进行无规随机游走, 即从任意节点出发, 以等概率游走到其邻居节点中的一个节点. 以图 2 为例, 若在某一时刻随机游走者处于节点 k 的位置, 那么下一时刻将分别以 $\frac{1}{4}$ 的概率游走到其邻居节点 $a, k-1, k+1$ 及 b , 其中 a 与 b 为节点 k 的长程连接邻居节点. 进一步, 我们将节点 1 设置为陷阱点, 考察从网络上其他节点到达陷阱节点 1 的平均首达时间. 我们知道在给定参数 c 的条件下, 随着幂指数 α 不断增加, 长程连接的长度越长, 出现的概率越小, 由于存在总能量约束, 因此长程连接的数量将会增加. 为了研究长程连接的数量和长度如何影响该空间网络上的随机游走过程, 我们首先进行了蒙特卡罗数值模拟. 当网络规模为 $N = 200, c = 1$ 时, 模拟结果显示, 一维空间网络上有陷阱的随机游走的平均首达时间与网络中添加长程连接的幂指数 α 存在如图 3 所示的关系, 图中每一个点是对 10^6 个网络样本上分别进行 5 次随机游走的平均结果. 可以发现存在一个最优的幂指数 α_0 在 1 附近使得网络上其他节点到陷阱节点的平均首达时间最短.

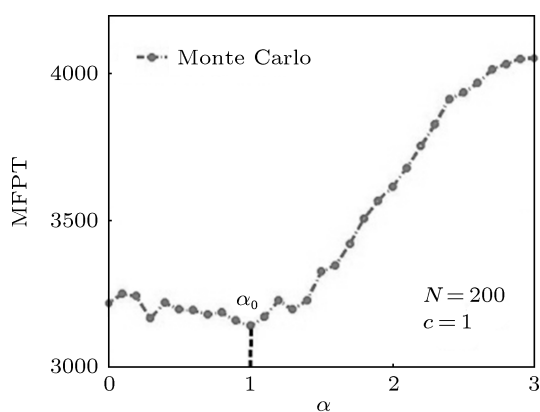


图 3 网络规模为 $N = 200$ 时的平均首达时间

2.3 平均首达时间的数学求解

事实上, 给定网络的邻接矩阵, 可以使用矩阵计算方式对陷阱问题进行理论分析. 定义网络的邻接矩阵为 A , 同时 T_{ij} 为随机游走者从网络上节点 i 首次到达节点 j 所需的时间. 由于网络节点的同质性, 我们只需计算以某个节点为陷阱节点, 其他所有节点首次到达该陷阱节点所需时间. 在规模为 N 的网络中, 不妨以节点 1 为陷阱点,

令 $\{T_{i1}\}, i \in \{1, 2, 3, \dots, N\}$ 表示从 i 节点首次返回陷阱节点 1 所需时间.

根据最小解理论和随机过程理论, 我们可以知道 $\{T_{i1}\}$ 是满足方程 $x_i = \sum_{k \neq 1, k \in \Omega} p_{ik} x_k + 1, i \in \Omega = \{1, 2, 3, \dots, N\}$ 的最小解, 其中 p_{ik} 是节点 i 经过一步游走到达节点 k 的概率, 是转移概率矩阵 $\tilde{P} = \text{diag}\{A\mathbf{1}\}^{-1}A$ 的第 i 行第 k 列元素. 如果将节点 1 看作陷阱点, 方程可简化为如下形式:

$$T_{i1} = \sum_{k \in \Omega_0} p_{ik} T_{k1} + 1, \quad i \in \Omega_0 = \{2, 3, \dots, N\}.$$

$$\text{若记 } \mathbf{T} = \begin{pmatrix} T_{21} \\ \vdots \\ T_{N1} \end{pmatrix}, \quad P = \begin{pmatrix} p_{22} & \cdots & p_{2N} \\ \vdots & \ddots & \vdots \\ p_{N2} & \cdots & p_{NN} \end{pmatrix},$$

$$\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{N-1}, \quad \text{可得 } \mathbf{T} = P\mathbf{T} + \mathbf{1}.$$

不难发现, P 是转移概率矩阵 \tilde{P} 去掉第 1 行第 1 列后得到的矩阵. 进一步可以得到 $\mathbf{T} = (I_{N-1} - P)^{-1}\mathbf{1}$, \mathbf{T} 的第 m 行元素代表 $m+1$ 节点首次返回陷阱节点 1 的时间, \mathbf{T} 的均值即为我们讨论的陷阱问题的平均首达时间.

2.4 数值模拟

基于上述理论分析, 只要能够得到前述构建的一维有限能量约束下的空间网络的状态转移矩阵, 就可以求解陷阱问题的精确结果. 考虑到该状态转移矩阵的复杂性, 我们使用计算机进行了相应的数值分析. 图 4 分别给出了

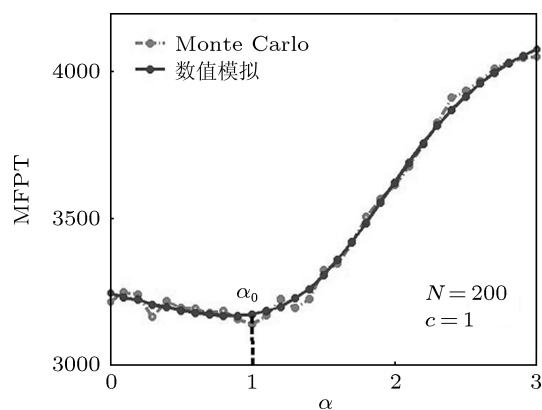


图 4 蒙特卡洛模拟结果与基于理论分析的数值模拟结果

网络规模 $N = 200$, $c = 1$ 时, 根据理论分析公式数值解出的陷阱问题的平均首达时间以及直接使用蒙特卡罗模拟方法得到的陷阱问题的平均首达时间, 其中数值解是在同样参数条件下对 4×10^5 个随机网络的平均值. 不难看出, 基于理论分析的数值模拟结果和直接使用蒙特卡罗方法得到的模拟结果能够很好地吻合, 而且相比于直接使用蒙特卡罗方法, 基于理论分析的数值模拟结果更加光滑、准确, 且耗时更短.

3 网络空间结构对平均首达时间的影响

3.1 网络规模对平均首达时间的影响

为了讨论网络空间结构对陷阱问题的平均首达时间的影响, 基于前述理论分析结果, 我们模拟了不同网络规模 N 下的陷阱问题. 取 $\alpha = 1$ 且 $c = 1$, 模拟结果如图 5 所示, 不难发现, 当网络规模 N 逐渐增大时, 平均首达时间 MFPT 与网络规模 N 呈幂律增长模式.

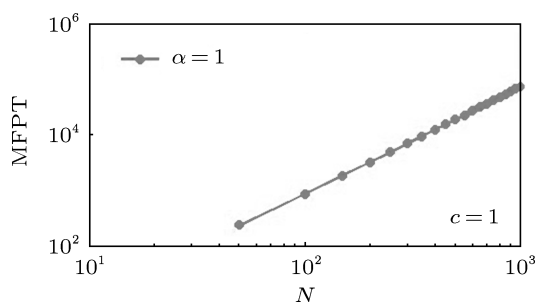


图 5 平均首达时间与网络规模的双对数坐标曲线

当 $c = 1$ 时, 我们考察了其他网络结构幂指数 α 下, 平均首达时间 MFPT 与网络规模 N 之间的关系, 结果发现网络规模 N 从 50 变化到 1000 时, 平均首达时间 MFPT 与网络规模 N 两者之间都存在很好的幂率关系, 不妨令 β 表示其幂指数, 即令 $\text{MFPT} \propto N^\beta (\beta > 0)$. 考察网络参数 α 和幂指数 β 之间的关系, 我们得到了如图 6 所示结果. 从图中不难看出, 当参数 α 从 0 变化到 3 时, 幂指数 β 随 α 的增加先减后增, 在 $\alpha = 1.5$ 的附近取得最小值. 这一结果表明, $\alpha = 1.5$ 时, 平均首达时间随网络规模增加的速率最慢. 根据此结果, 我们可以合理推测, 当网络规模 N 足够大时, 使得平均首达时

间 MFPT 取最小值的最优幂指数 α_0 可能在 1.5 左右. 这一推论与下面对网络规模对最优幂指数的讨论相符.

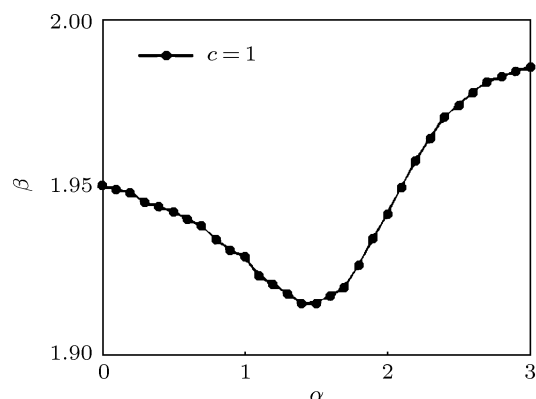


图 6 网络幂指数 α 与平均首达时间幂指数 β 之间的关系

3.2 总能量约束对平均首达时间的影响

由于我们构建的空间网络受到总能量 A 的约束, 在固定幂指数 α 的情况下, 增加总能量约束将会使得整个网络中的长程连接数量增加, 进而会影响相应的状态转移矩阵, 从而影响陷阱问题的平均首达时间. 由于总能量约束满足 $A = cN$, 当固定网络规模 $N = 200$ 时, 改变参数 c 就能够改变整个网络的总能量. 如图 6 所示的模拟结果显示, 固定幂指数 $\alpha = 1$, 当 c 增加时, 陷阱问题的首次到达时间将会减小.

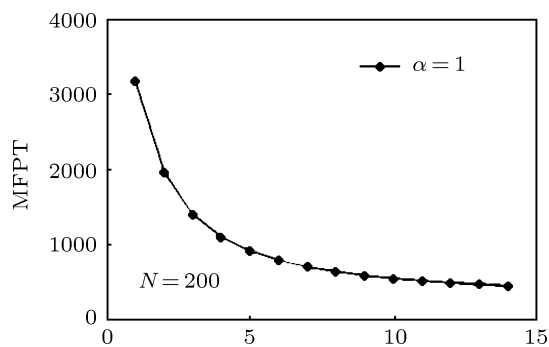


图 7 网络规模为 $N = 200$ 时平均首达时间与总能量约束的关系

4 网络空间结构对最优幂指数的影响

4.1 网络规模对最优幂指数的影响

空间网络考虑了网络节点的空间地理位置因素的影响, 因而随着网络空间结构性质的改变, 网

路上的随机游走过程也可能会发生重大改变. 显然, 随着网络规模的增加, 我们构建的空间网络结构会发生相应的变化. 为了更加深入研究网络规模对空间网络上的陷阱问题的影响, 我们进一步讨论了网络规模 N 对陷阱问题的最优幂指数 α_0 的影响. 基于前文的理论分析结果, 我们模拟了不同网络规模下的陷阱问题的平均首达时间. 为了更好地观察, 我们对模拟结果的数据进行了一定的数学处理, 旨在保留变化趋势, 去除规模效应. 具体做法是: 网络规模为 N 时, 当幂指数 α 从 0 变化到 3 时 (间隔 0.1), 对于每个指数 α 我们都将得到一个平均首达时间, 进而可以计算其平均值. 我们将原始的平均首达时间用该平均值重新标度 (除以该平均值), 就可以得到图 8 中对应于该网络规模 N 的一条曲线. 不同规模下重新标度后的具体模拟结果如图 8 所示, 随着 N 的增大, 最优幂指数 α_0 将会增加, 并且有稳定在某一特殊值的趋势. 结合前文我们所讨论的网络幂指数 α 与平均首达时间幂指数 β 之间的关系, 我们可以推测当网络规模 N 足够大时, 最优幂指数 α_0 可能稳定在 1.5 附近.

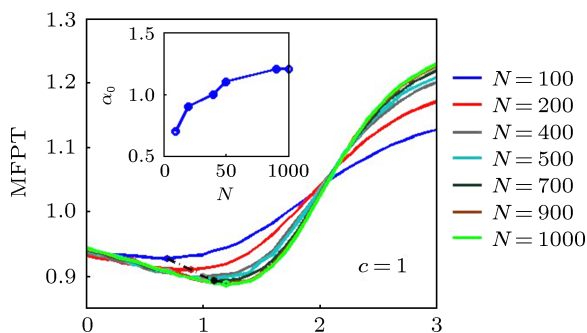


图 8 网络规模 N 与最优幂指数 α_0 的关系

4.2 总能量约束对最优幂指数的影响

为了讨论总能量约束可能对陷阱问题的影响, 我们也模拟了不同总能量约束下陷阱问题的平均首达时间. 同样地, 我们对不同总能量约束即参数 c

下的陷阱问题平均首达时间的数据也进行了前文所述的处理. 具体模拟结果参见图 9, 当固定网络规模为 $N = 200$ 时, 随着总能量约束增加, 最优幂指数 α_0 也相应增加, 并且也趋于稳定.

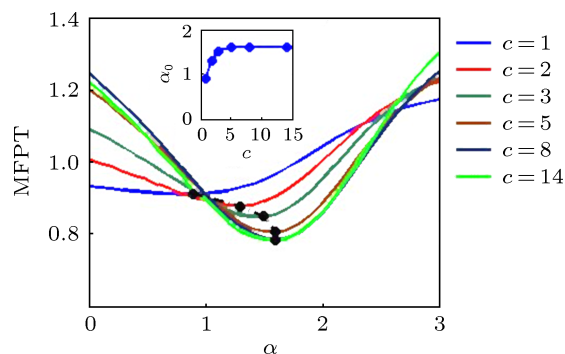


图 9 总能量约束与最优幂指数 α_0 的关系

5 结论

本文首先建立了一维有限能量约束下的空间网络, 并在此一维空间网络上讨论了随机游走过程中的陷阱问题. 我们发现存在一个最优的幂指数 α_0 使得该空间网络上的陷阱问题平均首达时间最短. 同时, 我们也从网络规模以及总能量约束两个角度分别讨论了空间网络结构对该最优幂指数 α_0 的影响. 研究结果显示, 陷阱问题的平均首达时间 MFPT 与网络规模 N 之间满足幂率关系, 不妨假设其幂指数为 β . 通过研究网络幂指数 α 与该幂指数 β 之间的关系, 我们推断当网络规模 N 足够大的时候, 使得平均首达时间 MFPT 取得最小值的最优幂指数 α_0 最终可能稳定在 1.5 附近. 我们的工作一方面丰富了网络上的随机过程的研究内容, 另一方面也加深了对空间网络的认识. 在这一研究方向上可待研究的问题还有很多, 例如精确求解我们发现的最优幂指数 α_0 、高维空间网络上的随机行走等问题. 我们需要进行更深入地研究, 以期更好地理解空间网络上的随机游走过程.

[1] Xu D, Li X, Wang X F 2007 *Acta Phys. Sin.* **56** 1313 (in Chinese) [许丹, 李翔, 汪小帆 2007 物理学报 **56** 1313]
 [2] Boccaletti S, Latorab V, Morenod Y, Chavez M, Hwang D U 2006 *Physics Reports* **424** 175

[3] Weng W G, Ni S J, Shen S F, Yuan H Y 2007 *Acta Phys. Sin.* **56** 1938 (in Chinese) [翁文国, 倪顺江, 申世飞, 袁宏永 2007 物理学报 **56** 1938]
 [4] Noh JD, Rieger H 2004 *Phys. Rev. Lett.* **92** 118701

- [5] Fronczak A, Fronczak P 2009 *Phys. Rev. E* **80** 016107
- [6] Herrero C P 2005 *Phys. Rev. E* **71** 016103
- [7] Campos P R A, Moreira F G B 2005 *Phys. Rev. E* **71** 061921
- [8] Wang S P, Pei W J 2009 *Physica A* **388** 514
- [9] Masuda N, Konno N 2004 *Phys. Rev. E* **69** 066113
- [10] Lee S, Yook S H, Kima Y 2008 *Physica A* **387** 3033
- [11] Aldous D 1990 *SIAM Journal on Discrete Mathematics* **3** 450
- [12] Tadi B, Rodgers G J, Thurner S 2007 *International Journal of Bifurcation and Chaos* **17** 2363
- [13] Kleinberg J M 2000 *Nature* **406** 845
- [14] Cajueiro D O 2009 *Phys. Rev. E* **79** 046103
- [15] Kim B J, Yoon C N, Han S K, Jeong H 2002 *Phys. Rev. E* **65** 027103
- [16] Zhou J, Liu Z H 2009 *Physica A* **388** 1228
- [17] Jiang Z H, Wang H, Gao C 2011 *Acta Phys. Sin.* **60** 058903 (in Chinese) [姜志宏, 王晖, 高超 2011 物理学报 **60** 058903]
- [18] Zhang Z Z, Qi Y, Zhou S G, Gao S Y, Guan J H 2010 *Phys. Rev. E* **81** 016114
- [19] Zhang Z Z, Lin Y, Gao S Y, Zhou S G, Guan J H, Li M 2009 *Phys. Rev. E* **80** 051120
- [20] Zhang Z Z, Wu B, Zhang H J, Zhou S G, Guan J H, Wang Z G 2010 *Phys. Rev. E* **81** 031118
- [21] Haynes C P, Roberts A P 2008 *Phys. Rev. E* **78** 041111
- [22] Tejedor V, Bénichou O, Voituriez R 2009 *Phys. Rev. E* **80** 065104
- [23] Agliari E 2008 *Phys. Rev. E* **77** 011128
- [24] Zhang Z Z, Qi Y, Zhou S G, Xie W L, Guan J H 2009 *Phys. Rev. E* **79** 021127
- [25] Zeng A, Zhou D, Hu Y Q, Fan Y, Di Z R 2011 *Physica A* **390** 3962

Random walks on spatial networks*

Dou Fei-Ling Hu Yan-Qing Li Yong Fan Ying Di Zeng-Ru†

(Department of Systems Science, School of Management and Center for Complexity Research, Beijing Normal University, Beijing 100875, China)

(Received 21 November 2011; revised manuscript received 20 February 2012)

Abstract

In this paper, we construct a cost constrained spatial network by adding long-range connections to the one-dimensional circle. The probability for a long-range connection between nodes i and j is $p_{ij} \propto d_{ij}^{-\alpha}$ ($\alpha \geq 0$), where d_{ij} is the lattice distance and the total length of the long-range connections is set to be $A = cN$ ($c \geq 0$), where c is a positive constant and N is the network size. According to the simulation and numeric results, we find an optimal power-law exponent α_0 such that the mean first-passage time is shortest. Furthermore, the shortest mean first-passage time has the power law relationships with the network size N . With the increase of network size N and the total cost A , the optimal power-law exponent α_0 increases monotonically and approaches 1.5.

Keywords: spatial networks, random walks, total energy constraint, mean first-passage time

PACS: 89.75.Hc, 89.75.Fb

* Project supported by the Fundamental Research Funds for the Central Universities and NSFC (Grants Nos. 60974084, 61174150) and NCET-09-0228.

† E-mail: zdi@bnu.edu.cn