

基于引力场理论的复杂网络路由选择策略研究*

刘刚[†] 李永树

(西南交通大学地球科学与环境工程学院, 成都 610031)

(2012年5月23日收到; 2012年6月20日收到修改稿)

利用引力场理论对网络传输过程中节点激发的引力场进行了描述, 建立了节点的引力场方程, 引入 α 和 γ 两个参数, 用于调节数据传输对节点畅通程度、节点传输能力和路径长度的依赖程度. 基于节点的引力场, 提出了一种高效的路由选择算法, 该算法下数据包将沿着所受路径引力最大的方向进行传递. 为检验算法的有效性, 引入有序状态参数 η , 利用其由自由流到拥塞态的指标流量相变值度量网络的吞吐量, 并通过节点的介中心值 B 分析网络的传输性能和拥塞分布. 针对算法在不同 α, γ 取值条件下的路由情况进行了仿真. 仿真结果显示, 与传统最短路由算法相比, 本文算法将网络传输能力提高了数倍, 有效地均衡了节点的介中心值分布, 传输路径平均长度 (L_{avg}) 随负载量 R 的增加表现出先增后减的变化趋势, 而参数 α 与 γ 值的变化对网络传输能力几乎没有影响, 说明本文路由算法的性能不依赖于 α 与 γ , 对于可行域内任意的 α 与 γ , 算法都能保证网络传输能力近似相等.

关键词: 复杂网络, 引力场, 路由策略, 拥塞

PACS: 89.75.Hc, 89.20.Hh

1 引言

复杂网络理论是近年来系统科学和复杂性科学领域的前沿热点问题之一, 已经吸引了大量来自物理、数学、生物、计算机、社会学等领域的学者对此进行深入研究^[1-4]. 目前, 复杂网络理论研究主要集中在网络拓扑结构复杂性^[5-7]及网络动力学过程^[8-10]两个方面. 在网络拓扑结构复杂性方面, 学者们希望通过对现实系统结构复杂性的研究发现系统的演化机理, 并为提高网络上信息传输效率探索最优的网络拓扑结构. 对于某些基础设施网络如道路交通网络、电力网络等, 网络结构复杂且覆盖范围非常广, 如果从整体上改变其拓扑结构, 势必会给周围环境带来较大影响, 且经济代价太大. 所以, 通过调节网络上的动力学行为以提高网络传输效率已经成为复杂网络理论研究中主要课题之一.

拥塞是许多真实网络普遍存在的一种复杂现

象, 而解决拥塞的主要方法之一是通过设计高效的路由选择策略以提高网络容量并缓解其拥塞程度^[11]. 在有关网络路由策略的研究中, Noh 等^[12]分析了复杂网络上的随机行走过程并建立了两个节点之间平均首次通过时间的精确表达形式, 发现走向聚类系数较大节点的随机行走比走向其他节点更快, 且在给定的时间间隔内访问聚类系数较大的节点的随机行人更多. Ramascc 等^[13]利用极值优化思想研究了路径长度限制下的路由选择策略, 在一定程度上提高了路由策略的可扩展性. 已有研究表明, 网络的传输能力与节点的最大介数之间存在近似反比关系^[14], 当网络规模确定时, 最大介数越大, 则网络传输能力越小. 受该论点的启发, 学者们开始以实现“节点最大介数最小化”思想进行网络路由策略的优化研究. Danila 等^[15]针对复杂网络上的交通优化问题提出了一种启发式路由算法, 该算法通过节点最大介数的最小化实现网络交通流的均衡化, 与最短路由算法相比, 该算法在网络处于低负载时其平均传输时间略有增加, 但较大地

* 高等学校博士学科点专项科研基金 (批准号: 20100184110019)、2013 年西南交通大学博士研究生创新基金和中央高校基本科研业务费专项资金资助的课题.

[†] E-mail: liuganggis@sina.com

降低了网络在拥塞状态下数据包的平均传输时间. 在 Danila 等给出的算法基础上, Kawamoto 等^[16]提出了一种改进的启发式路由算法, 通过减少最大介数最小化的迭代次数进而降低交通流均衡化计算的收敛时间.

从数据传输过程可以看出, 数据包往往汇聚于介数较大的节点处进而引发拥塞. 其实, 我们可以从引力理论的角度分析这一现象, 即介数较大的节点对网络上大量数据包的传递具有更大的吸引力. 利用引力理论, 钱江海等^[17]通过建立空间网络的引力模型研究了节点之间的潜在预期流; 刘刚等^[18]以数据包传递的邻居节点及邻居节点到目标节点的最短路径上最拥塞的节点来刻画该最短路径对数据包的引力作用, 并提出了一种基于引力约束的路由选择策略, 在一定程度上提高了网络的吞吐量并缓解了网络拥塞情况. 节点作为数据包的载体直接影响其网络传输的动态机理, 且这种机理可以表现为节点对数据包的引力作用. 在不同的路由选择算法中, 节点的引力作用不同. 以最短路由算法为例, 大部分数据包的传递都会经过度较大的节点, 说明此类节点具有更大的吸引力. 如果路由选择过程不再像最短路由算法那样具有确定性, 那么要保证网络具有较高的传输效率, 就需要考虑网络上节点的拥塞状态, 尽量选择一条较为通畅的路径进行传递. 此时, 单个节点的引力作用不再占绝对优势, 我们可以将节点的引力作用及节点之间的相互作用转化为传输路径对数据包的引力作用, 认为较畅通且长度较短的路径具有更大的吸引力.

为进一步研究节点对数据包传递的影响, 本文利用引力场理论对网络传输过程中节点激发的引力场进行了描述, 并建立了节点的引力场方程及传输路径对数据包的引力计算公式. 在此基础上, 提出一种基于节点引力场的复杂网络路由选择策略, 该策略顾及了传输路径的长度、路径上节点的拥塞程度及节点传输能力等因素的影响. 我们预测, 与最短路由算法相比, 本文路由策略将有效地均衡网络负载、缓解网络拥塞情况并提高网络的吞吐量.

2 基于引力场的复杂网络路由策略

2.1 节点的引力场描述

在传统的最短路由算法中, 数据包总是沿着距

离目标节点最近的路径进行传递, 只要源节点和目标节点确定, 则传输路径已确定. 文献 [18] 指出可以利用引力理论来研究网络上数据包的传输过程并提出了一种具有引力约束的路由算法. 节点的介数在很大程度上反映了节点对数据包传递过程的引力作用. 介数越大, 则节点对数据包的引力就越大, 就会吸引更多的数据包经过该节点, 从而也易导致交通拥塞的发生. 拥塞的后果集中表现为某些节点的缓存队列中有大量数据包等待传输, 而另一些节点则可能处于空闲状态. 在最短路由算法下, 节点的引力作用是可以确定的, 且引力分布极不均匀. 从逆向思维的角度, 如果能有效地控制节点的引力作用以消减引力分布的不均匀性, 使数据包在传输过程中能选择性地避开某些拥塞严重的节点, 或许能在一定程度上提高网络的传输能力并改善其拥塞情况.

分析最短路由策略不难得出, 介数较大的节点吸引大量的数据包汇聚于此进而引发拥塞, 且缓存队列中数据包的个数既反映了节点的拥塞程度, 也反映了其引力大小. 所以, 如果要改善节点的引力分布, 须以节点缓存数据包个数及节点传输能力为主要考察因素, 通过及时更新数据包的传输路径以实现网络传输过程中节点引力的动态分配.

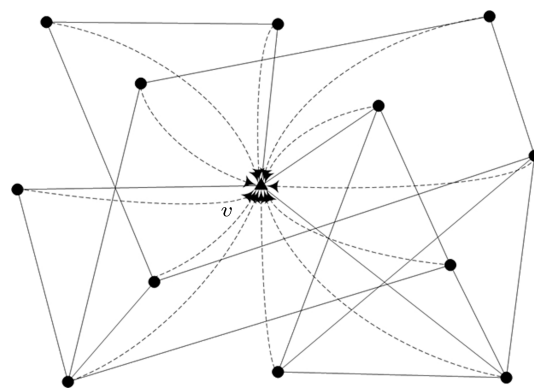


图1 节点的引力场

引力场理论最先由麦克斯韦提出并用于描述电磁现象, 后由爱因斯坦将其发展为广义相对论. 在宇宙空间中, 引力场是暗能量和星体相互作用的产物, 引力场中某一点的引力与暗能量的虚拟质量和星体的质量的乘积成正比, 与该点到旋转中心的距离的平方成反比, 且引力与物体的质量无关. 根据引力场理论, 我们可以对网络上数据包的传输过程做如下抽象: 数据包总是同时受到网络上不同节

点的引力, 对于任意节点的引力, 数据包在一定程度上都有流向该节点的趋势, 且节点对数据包的引力作用与该节点的传输能力及拥塞程度、数据包与节点之间的逻辑距离等因素有关, 而与数据包自身无关. 这样, 每一个节点对网络上所有数据包都存在一定的引力作用, 由此可以激发一个引力场, 而数据包在不同节点的引力场作用下如何传递, 将取决于具体的路由选择策略, 且路由选择策略主要考虑如何进行引力场的控制及基于引力场进行路径选择. 在不考虑数据包传输的目标节点的前提下, 假设数据包当前所在节点 j , 则某一节点 i 对该数据包的引力为

$$F_i = \begin{cases} \kappa \frac{c_i^{2\alpha}}{d_{ij}^\gamma}, & q_i = 0, \\ \kappa \frac{(c_i/q_i)^\alpha c_i^\alpha}{d_{ij}^\gamma}, & q_i > 0, \end{cases} \quad (1)$$

上式可以看作是节点的引力场方程, 其中, F_i 为节点 i 对数据包的引力; κ 为常数; c_i 为节点 i 的传输能力, 即单位时间内所能处理的最大数据包个数; q_i 为节点 i 当前缓存队列中的数据包个数; c_i/q_i 可以看作当前节点 i 的畅通程度; d_{ij} 为节点 i 到节点 j 的最短路径长度; α 和 γ 为两个可调节参数, 分别用于调节数据传输对节点畅通程度、节点传输能力和路径长度的依赖程度, 且 $\alpha > 0, \gamma > 0$. 在引力场理论中, 引力场中某一点所受引力与暗能量的虚拟质量和星体的质量的乘积成正比, 与该点到旋转中心的距离的平方成反比, 且与物体的质量无关. 对于复杂网络而言, 节点等价于星体, 而节点的传输能力 c_i 等价于星体的质量, 传输能力越大, 则引力就越大; c_i/q_i 反映了节点 i 的畅通程度 (或相对通行能力), 等价于暗能量的虚拟质量, c_i/q_i 越大, 则节点对数据包的引力就越大; 数据包到节点的最短路径长度 d_{ij} 等价于某点到星体旋转中心的距离, d_{ij} 越大, 则引力越小. 由该引力场方程可以看出, 节点对数据包的引力作用与节点的传输能力和节点的畅通程度的乘积的 α 次幂成正比, 而与节点到数据包的最短路径长度的 γ 次幂成反比. 由此可见, (1) 式表达了与引力场理论相似的物理意义.

2.2 路由选择策略

对于给定的复杂网络, 假设网络上所有节点都具有路由、收包、发包的功能, 且初始状态网络负载为 0, 之后每一时间步产生 R 个数据包并随机地选择源节点和目标节点, 且产生的数据包自动添加

到源节点的缓存队列尾部, 在单位时间步内每个节点最多只能发送 c_i 个数据包, 节点的缓存队列长度无限且采用先进先出方式. 在网络传输过程中, 数据包总是由当前节点发送给某个邻居节点, 若该邻居节点为数据包的目标节点, 则删除该数据包; 否则, 按照给定的路由策略进入该邻居节点的缓存队列.

如图 2 所示, 假设数据包传递的源节点 s 、目标节点 t , 当前所在节点 v , 其邻居节点集为 N_v , 邻居节点 $i (i \in N_v)$ 到目标节点 t 的最短路径所经过的节点集为 N_{it} , 且设 N_{it} 包含节点 i 但不包含目标节点 t , 因为 t 为数据包传递的目标节点, 其交通状态对节点 v 处数据包的传输效率没有影响, 而其他节点的拥塞情况将直接影响 v 处数据包的传递. 由于 N_{it} 中每一个节点都会激发一个引力场, 即每一个节点都对数据包都存在引力作用. 在一定程度上, 数据包有可能沿着节点 $i (i \in N_v)$ 到目标节点 t 的最短路径传输, 但这种可能性并不是由路径上某一个节点所决定, 而是路径上所有节点共同作用的结果. 为此, 这里定义该最短路径对数据包的引力表达为路径上所有节点对数据包的引力的平均值

$$F_{it} = \frac{1}{n} (F_1 + F_2 + \dots + F_n) = \frac{1}{n} \sum_{j \in N_{it}} F_j, \quad (2)$$

其中, F_{it} 为邻居节点 i 到目标节点 t 的最短路径对数据包的引力, n 为 N_{it} 的节点数.

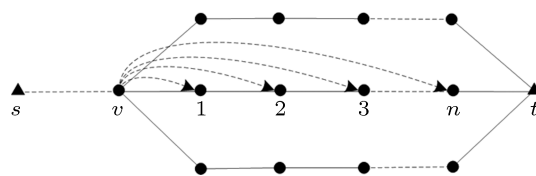


图 2 路由选择原理

从引力场理论的角度, 我们认为路径对数据包的引力 F_{it} 越大, 则数据包沿着该路径传递的概率就越大. 为此, 本文提出一种基于节点引力场的路由选择策略, 具体的路由选择过程如下:

- 1) 获取数据包当前所在节点 v 的邻居节点集 N_v .
- 2) 针对每一个邻居节点 $i (i \in N_v)$, 如果数据包的传输路径尚未经过节点 i , 获取 i 到目标节点 t 的最短路径, 并提取构成该路径的节点集 N_{it} , 计算 N_{it} 上所有节点对数据包的引力, 由此计算 i 到 t 的最短路径对数据包的引力 F_{it} .

3) 获取最短路径对数据包的引力最大时 (即 $F = \max\{F_{it}\}$) 所对应的邻居节点, 并以该邻居节点作为数据包传递的下一个路由节点. 如果存在多个引力相等且最大的节点, 则选择路径长度最短所对应的邻居节点, 若还存在多个符合条件的节点, 则随机地选择其中的一个节点作为下一个路由节点.

可见, 数据包在传递过程中总是选择引力最大的路径所对应的邻居节点作为下一个路由节点. 如图 3(a) 所示, 当前节点 v 存在两个邻居节点的最短路径可供选择, 两条路径分别经过节点 i, u, b 和节点 j, e, f, g , 当参数 $\alpha = 1$, $\gamma = 2$ 且节点传输能力 $c = 1$ 时, 路径引力分别为 $F_{it} = 1/3(1/2 + 1/16 + 1/27) = 0.1998$ 和 $F_{jt} = 1/4(1/2 + 1/16 + 1/27 + 1/32) = 0.1577$, 故选择节点 i 作为数据包传递的下一个路由节点; 当节点拥塞情况发生变化时, 如图 3(b) 所示, 经过节点 i, u, b 和节点 j, e, f, g 的两条路径对数据包的引力分别为 0.1998 和 0.342, 故将选择节点 j 作为下一个路由节点. 本文路由算法保证数据包始终沿着引力最大的路径进行传递, 且该路径是所有邻居节点对应的最短路径中较为畅通的路径.

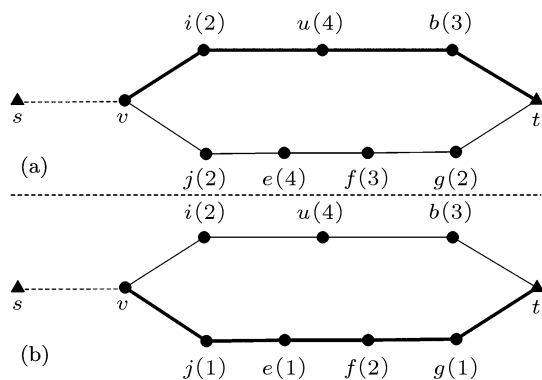


图3 路由选择示意图 (i, j, u 等表示节点 ID; () 中的数字表示节点缓存数据包个数; 加粗线条表示本文算法所选择的路径)

2.3 评价指标

2.3.1 有序状态参数

随着单位时间步内新增负载量 R 的增大, 网络交通状态将由自由流状态进入拥塞状态, 且存在一个临界负载量 $R_c^{[19]}$. 为分析网络交通流情况, 我们定义了一个拥塞指标, 用于描述网络交通状态由自由流到拥塞流的相变:

$$\eta(R) = \lim_{t \rightarrow \infty} \frac{1}{R} \frac{W(t)}{t}, \quad (3)$$

其中, $W(t)$ 是 t 时刻网络中的数据包个数; R 是单位时间步网络上新产生的数据包个数; $\eta(R)$ 表示截止 t 时刻网络上剩余数据包个数与总共产生的数据包个数的比值, 反映了网络处理数据包的能力. 当 $R < R_c$ 时, 单位时间步内网络上新增数据包个数与被处理完的数据包个数相等, 网络中剩余数据包个数 $W(t) = 0$, 拥塞指标 $\eta(R) = 0$; 当 $R > R_c$ 时, 单位时间步内网络上新增数据包个数明显大于被处理完的数据包个数, 网络迅速进入拥塞状态, $\eta(R) > 0$ 且变化非常突然. 所以, 使 $\eta(R) = 0$ 的最大负载量 R 就是网络的吞吐量 R_c .

2.3.2 节点介中心

中心性是度量复杂网络节点重要性的主要方法, 通过对网络中心性的量化分析可以发现网络上部分节点比其他节点更为重要^[20]. 节点介中心 (betweenness centrality, BC) 是中心性度量的一种方法, 它给出了在网络中一个节点处通过的最短路径, 反映了节点连通能力的聚集度^[21,22]. 为分析不同路由选择策略下网络中心性的变化, 这里将节点介中心值的定义做了适当修改: 定义节点介中心值为在确定路由选择策略下通过某一个节点的路径数量占所有路径数量的比值. 理论上, 不同路由选择策略下节点的介中心值是不同的. 为此, 我们设 $b_i^{(s,t)}$ 是以 s 为源节点、 t 为目标节点且在传输过程中经过节点 i 的数据包的个数, 则系统在 $[t_0, T + t_0]$ 时间内所有被处理完且经过节点 i 的数据包的个数为

$$b_i = \sum_{t_0}^{T+t_0} \sum_{s,t=1, s \neq t}^N b_i^{(s,t)}, \quad (4)$$

由此, 定义节点 i 的介中心值为

$$B_i = \frac{b_i}{\sum_{j=1}^N b_j}, \quad (5)$$

其中, B_i 为节点 i 的介中心值, N 为网络的节点数. B_i 反映了节点 i 连通能力的聚集度, B_i 越大, 则经过节点 i 的数据包越多, 越容易引发拥塞.

3 数值模拟及分析

为检验本文算法的有效性, 我们选择 BA 无标度网络模型^[23], 节点数 $N = 100$, 且 $m_0 = m = 4$. 在仿真过程中, 每次实现都运行 10000 步, 此时可认为网络上的流量基本处于稳态, 并根据最后 1000

步的平均值确定网络的吞吐量 R_c 和节点的介中心值 B_i , 并分别针对 α 和 γ 在不同取值下的网络路由效率进行对比实验. 此外, 数据包的最大传输时间 $\langle T_{\max} \rangle$ 、平均传输时间 $\langle T_{\text{avg}} \rangle$ 及传输路径的平均长度 $\langle L_{\text{avg}} \rangle$ 也是衡量网络传输性能的重要指标 [24,25], 为进一步评价路由算法的效率, 本文还将对 $\langle T_{\max} \rangle$ 、 $\langle T_{\text{avg}} \rangle$ 和 $\langle L_{\text{avg}} \rangle$ 进行统计分析. 为不失一般性, 我们将分别讨论节点传输能力等于常数 ($c = 1$) 和等于度 ($c = k$) 两种情况下网络交通流的动态特性.

图 4 表示最短路由算法在节点传输能力 $c = 1$ 时有序状态参数 η 、最大传输时间 $\langle T_{\max} \rangle$ 、平均传输时间 $\langle T_{\text{avg}} \rangle$ 及路径平均长度 $\langle L_{\text{avg}} \rangle$ 与负载量 R 的关系. 仿真结果显示, 当 $R = 5$ 时, η 发生相变, $\langle T_{\max} \rangle$ 与 $\langle T_{\text{avg}} \rangle$ 也突然增大, 说明最短路由算法下网络吞吐量 $R_c = 5$. 随着负载量 R 的增加, 路径平均长度 $\langle L_{\text{avg}} \rangle$ 逐渐减小, 其原因是随着 R 的增加, 网络越来越堵塞, 传输路径较长的数据包愈发不容易到达目标节点, 从而导致路径的平均长度逐渐减小. 图 5 表示本文路由算法在节点传输能力 $c = 1$ 时有序状态参数 η 、最大传输时间 $\langle T_{\max} \rangle$ 、平均

传输时间 $\langle T_{\text{avg}} \rangle$ 及平均传输距离 $\langle L_{\text{avg}} \rangle$ 与负载量 R 的关系, 其中给出了 α 与 γ 在 5 种不同组合取值下的统计结果. 由仿真结果可以看出: 当 $R < R_c$ 时, 有序状态参数 η 接近于 0, 最大传输时间 $\langle T_{\max} \rangle$ 与平均传输时间 $\langle T_{\text{avg}} \rangle$ 均较稳定, 数据包的平均传输距离 $\langle L_{\text{avg}} \rangle$ 随 R 的增加而增大且呈现加速增长趋势; 当 $R > R_c$ 时, η , $\langle T_{\max} \rangle$, $\langle T_{\text{avg}} \rangle$ 均迅速增大, 且 η , $\langle T_{\max} \rangle$, $\langle T_{\text{avg}} \rangle$ 均在 $R = 16$ 时发生相变, 由此可知网络吞吐量 $R_c = 16$; $\langle L_{\text{avg}} \rangle$ 随 R 的增加逐渐减小且较为缓慢, $\langle L_{\text{avg}} \rangle$ 在 $R = R_c$ 处取得最大值. 同时, 由仿真结果我们可以发现一个重要现象, 在 α 与 γ 的 5 种不同组合取值下, η , $\langle T_{\max} \rangle$, $\langle T_{\text{avg}} \rangle$ 及 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的变化趋势相同且在同一 R 值时近似相等, 说明在节点传输能力 $c = 1$ 时网络传输效率对 α, γ 不敏感. 由本文算法原理可知, 当负载量 R 较小时, 网络处于自由流状态, 路由选择过程类似于最短路由算法, 数据包将主要沿着最短路径进行传递; 随着 R 的增大且 $R < R_c$ 时, 网络上部分节点的缓存队列开始汇集少量数据包, 这时数据包将选择较为畅通的路径进行传递, 避让那些产生堵塞的节点, 从而导致传输路径的平均长度增

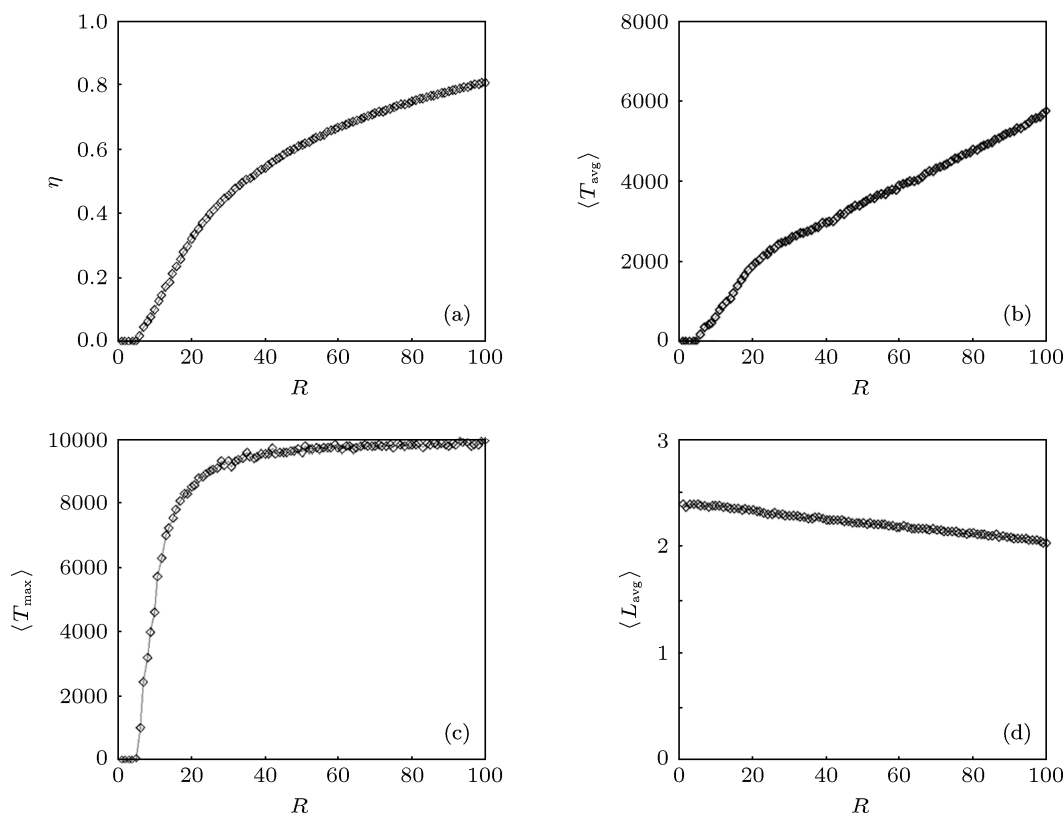


图 4 最短路由算法有序状态参数 η 、最大传输时间 $\langle T_{\max} \rangle$ 、平均传输时间 $\langle T_{\text{avg}} \rangle$ 及平均传输距离 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的变化情况 ($N = 100, m_0 = m = 4, c = 1$)

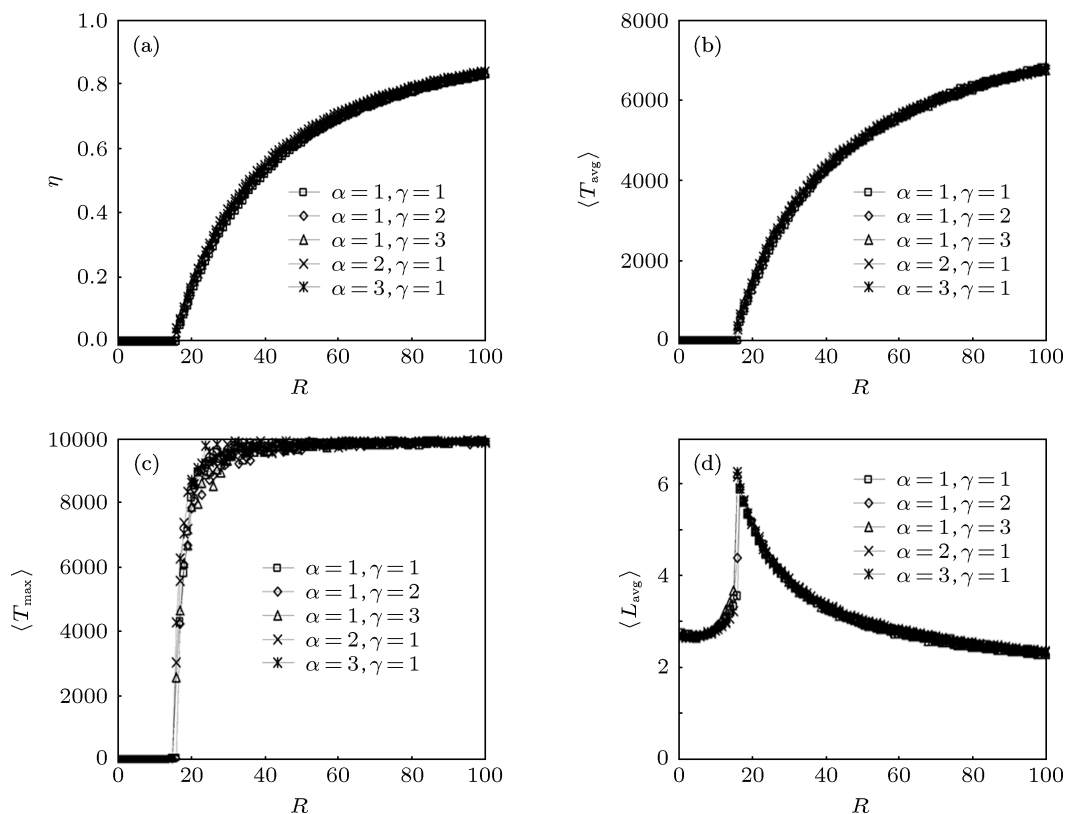


图5 本文路由算法有序状态参数 η 、最大传输时间 $\langle T_{\max} \rangle$ 、平均传输时间 $\langle T_{\text{avg}} \rangle$ 及平均传输距离 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的变化情况 ($N = 100, m_0 = m = 4, c = 1$)

大, 本文算法采用的这种避让策略缓解了已拥塞节点的拥塞程度并有效地控制了拥塞的恶化, 故整个网络不会陷入拥塞状态; 随着 R 的持续增大, 算法的避让策略将无法控制节点拥塞程度的加剧, 引发越来越多的节点陷入拥塞, 最终导致网络陷入拥塞状态, 此时数据包将不再避让某些节点, 传输路径的平均长度将逐渐减小。

图6表示最短路由算法在节点传输能力 $c = 1$ 时在不同负载量 R 下节点介中心值的分布情况, 其中给出了 $R = 1, 5, 20, 60$ 下的介中心值分布情况。仿真结果显示, 在 $R = 1$ 时节点介中心值分布最不均匀, 随着 R 的增加, 不均匀程度逐渐减弱但仍然存在明显的不均匀性。图7表示本文算法在节点传输能力 $c = 1$ 时在不同负载量 R 下节点介中心值的分布情况, 分别对 $R = 1, 10, 16, 20, 40, 60, 80, 100$ 等8种负载量下的节点介中心值进行了统计, 且 $\alpha = 1, \gamma = 2$ 。由仿真结果可以看出, 随着负载量 R 的增加, 节点介中心值分布表现出“不均匀分布—均匀分布—不均匀分布”的变化趋势; 当 $R = 1$ 时, 节点介中心值分布最不均匀, 且分布情况与最短路由策略相似, 随着 R 的增加, 介中心值分布逐渐趋于均匀, 并在 $R \approx 20$ 时均匀程度达

到最大, 之后介中心值分布又随 R 的增加变得不均匀; 当 $R = R_c$ 时, 节点介中心值分布已经均匀了很多, 但分布最均匀时对应的负载量 $R \approx 20$, 所有节点的介中心值几乎相等且近似等于 0.01。

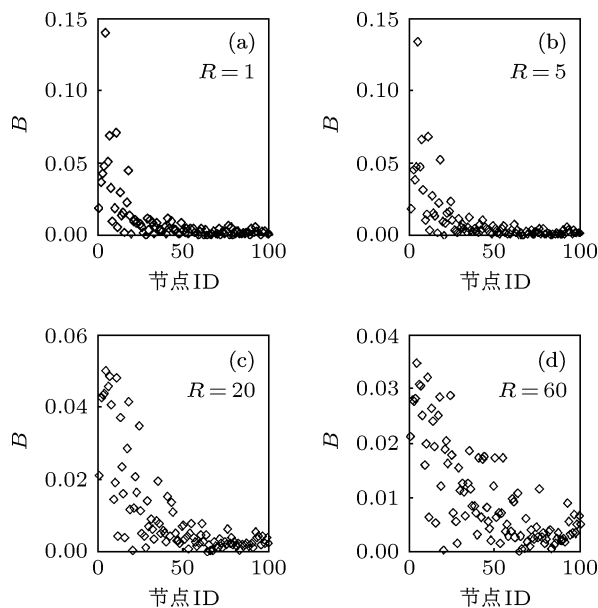


图6 最短路由策略在不同负载量 R 下节点的介中心值分布情况 ($N = 100, m_0 = m = 4, c = 1$)

我们对最短路由算法和本文路由算法下节点的拥塞情况进行了统计分析,如图 8 所示为在节点传输能力 $c = 1$ 时网络上所有节点的拥塞情况,即节点缓存队列中的数据包包数 $N(P)$. 图 8(a),(b) 均是在负载量 $R > R_c$ 条件下得到的结果,所以网络处于拥塞状态. 由仿真结果可以看出,最短路由策略下网络流量分布极不均匀,其中 58% 的节点拥塞程度为 0, 90% 的节点拥塞程度小于 3, 且最大拥塞为 25712; 本文路由算法下网络节点的流量分布较均匀, 度较大节点的拥塞情况得到了明显改善,

97% 的节点拥塞程度在 100 到 700 之间, 且最大拥塞为 703, 与最短路由算法相比, 最大拥塞程度降低了近 37 倍. 由本文路由算法原理可知, 数据包在传递过程中很好地考虑了路径上所有节点的拥塞程度, 从而可以及时、有效地避开拥塞较为严重的节点, 将数据包传递给其他较为空闲的节点, 逐步缓解拥塞较严重节点的拥挤程度, 最终实现整个网络负载的均衡化. 同时, 由于数据包的传递顾及了路径长度的影响, 所以数据包不会在网络上过多地迂回传递, 从而保证网络具有较强的传输能力.

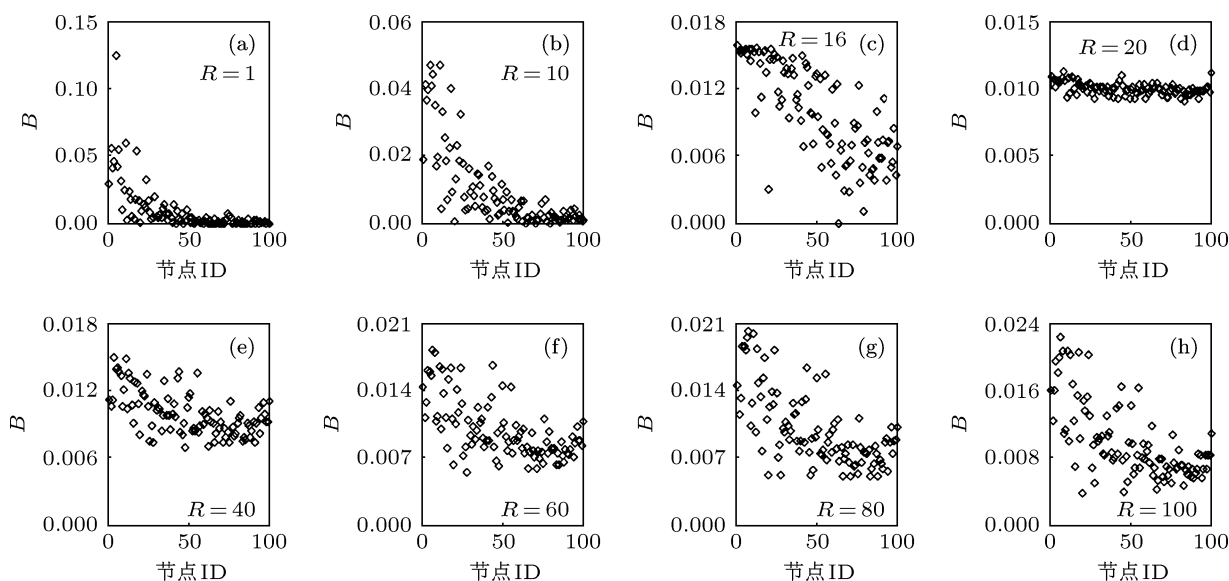


图 7 本文路由策略在不同负载量 R 下节点的介中心值分布情况 ($N = 100, m_0 = m = 4, c = 1, \alpha = 1, \gamma = 2$)

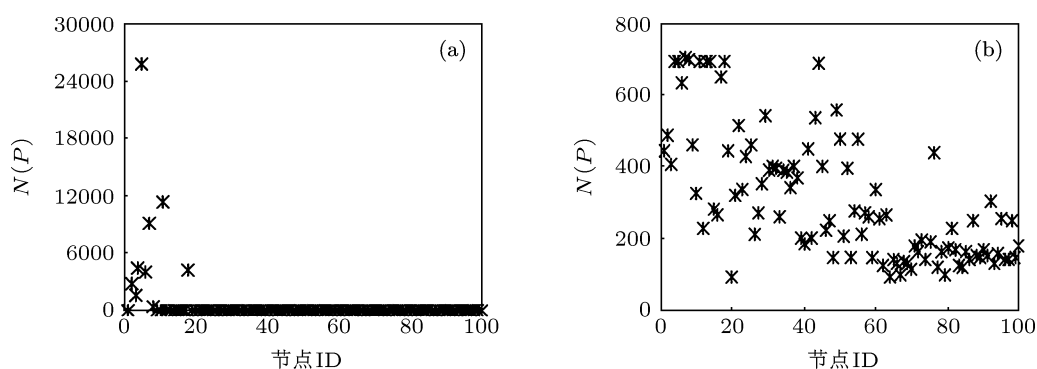


图 8 网络拥塞情况仿真结果 ($N = 100, m_0 = m = 4$) (a) 最短路由算法 ($R = 20, c = 1, N(P)$ 总量为 63716); (b) 本文路由算法 ($R = 20, c = 1, \alpha = 1, \gamma = 2, N(P)$ 总量为 30864)

为进一步分析本文算法的性能, 我们还针对节点传输能力等于度 ($c = k$) 时的网络交通流特性进行了实验, 图 9 为 $c = k$ 时得到的仿真结果. 仿真结果显示, 当 $c = k$ 时, 网络吞吐量 $R_c = 130$,

说明增大节点的数据处理能力可以有效地提高网络吞吐量. 当 $c = k$ 时, 参数 α 与 γ 值的变化对网络传输性能没有什么影响, 在 α 与 γ 的 5 种不同组合取值下, 有序状态参数 η 随负载量 R 的变

化几乎一样, 最大传输时间 $\langle T_{\max} \rangle$ 与平均传输时间 $\langle T_{\text{avg}} \rangle$ 也表现出类似特性. 由此可见, 对于节点传输能力 $c = 1$ 和 $c = k$ 两种情况, 网络传输过程均表现出“传输性能不依赖于 α 和 γ 值”这一重要动力学现象. 与 $c = 1$ 相比, 在 α 与 γ 的不同取值下, 路径平均长度 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的变化不尽相同. 从总体变化趋势上看, $\langle L_{\text{avg}} \rangle$ 仍表现出先增加后减小的趋势, 但当 $R < R_c$ 时, 参数 α 与 γ 值的变化对 $\langle L_{\text{avg}} \rangle$ 的影响较大. 针对不同的 α, γ 值, $\langle L_{\text{avg}} \rangle$ 的最大值不同且达到最大时对应的负载量 R 也不一样, 且均不在 $R = R_c$ 处. 当 $\alpha = 1, \gamma = 2$ 时, $\langle L_{\text{avg}} \rangle$ 在 $R \approx 90$ 时取得最大且 $\langle L_{\text{avg}} \rangle \approx 7.178$, 而当 $\alpha = 3, \gamma = 1$ 时, $\langle L_{\text{avg}} \rangle$

在 $R \approx 70$ 时取得最大且 $\langle L_{\text{avg}} \rangle \approx 9.782$. 同时可以看出, 在 $R < R_c$ 时, α 与 γ 表现出对 $\langle L_{\text{avg}} \rangle$ 不同的影响机理, 当 $\alpha = 1$ 时, 路径平均长度 $\langle L_{\text{avg}} \rangle$ 随 γ 值的增加而减小, 其原因是 γ 的增加导致增强了路径选择过程对路径长度的依赖程度, 使得数据包会尽量选择长度较短的路径进行传递; 当 $\gamma = 1$ 时, 路径平均长度 $\langle L_{\text{avg}} \rangle$ 随 α 的增加而增大, 其原因在于 α 的增加等价于增强路径选择过程对路径畅通度与节点传输能力的依赖程度, 于此同时也就降低了对路径长度的依赖程度, 从而导致路径平均长度变大; $\langle L_{\text{avg}} \rangle$ 在 $R = R_c$ 处实现一致, 且当 $R \geq R_c$ 时, 参数 α 与 γ 值的变化将对 $\langle L_{\text{avg}} \rangle$ 没有影响.

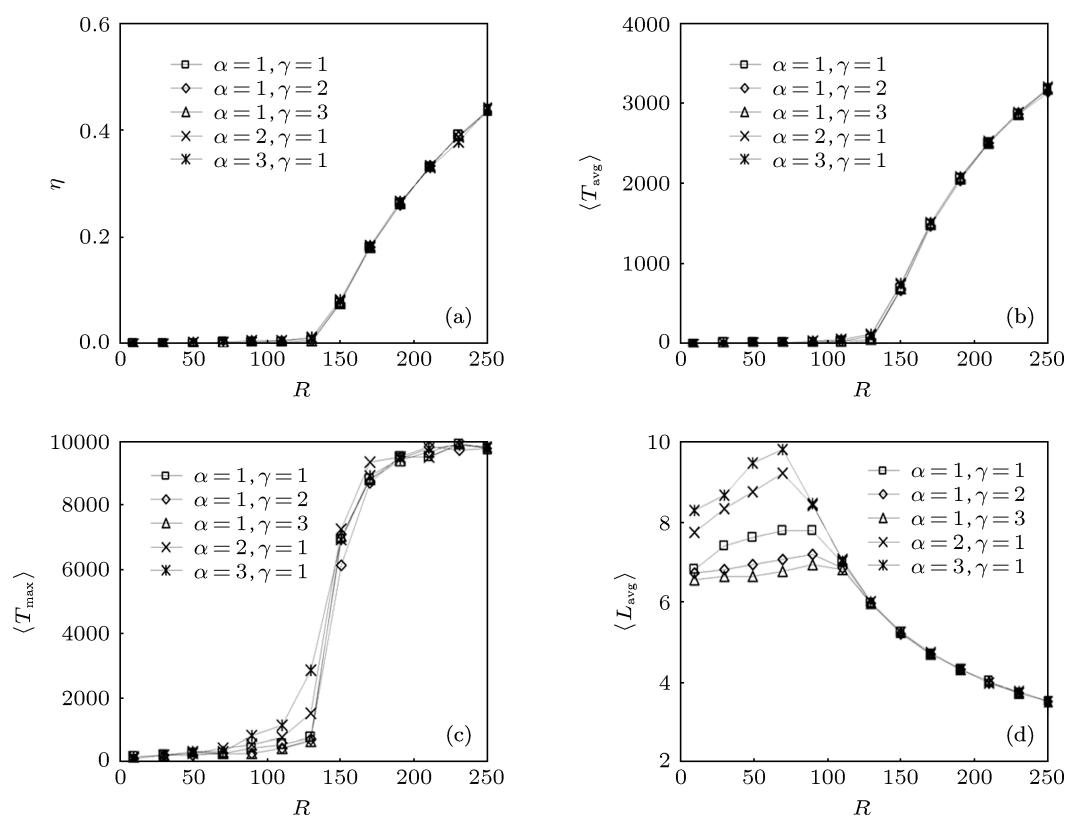


图9 本文路由算法有序状态参数 η 、最大传输时间 $\langle T_{\max} \rangle$ 、平均传输时间 $\langle T_{\text{avg}} \rangle$ 及平均传输距离 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的变化情况 ($N = 100, m_0 = m = 4, c = k$)

图 10 表示本文路由策略在节点传输能力 $c = k$ 时在不同负载量 R 下节点的介中心值分布情况. 结果显示, 节点介中心值分布的不均匀程度随负载量 R 的增加先减小后增加, 且在 $R \approx 150$ 时分布的非均匀性相对较弱, 几乎没有 $B = 0$ 的节点, 说明所有节点均得到了有效利用. 但与 $c = 1$ 时的节点介中心值分布相比, $c = k$ 时节点的介中心值

分布没有表现出 $c = 1$ 时的均匀分布特性. 分析其原因可知, 当 $c = k$ 时, 度较大节点的传输能力增大, 导致节点对数据包的引力增大, 进而使得经过该节点的路径对数据包的引力也增大, 这样就会有更多的数据包选择该路径进行传递, 导致度较大节点的介中心值增大, 但这并不是表明此类节点的堵塞程度很大, 就其与度较小节点的相对

拥塞情况而言 (q/c), 其拥塞程度又远小于最短路由策略.

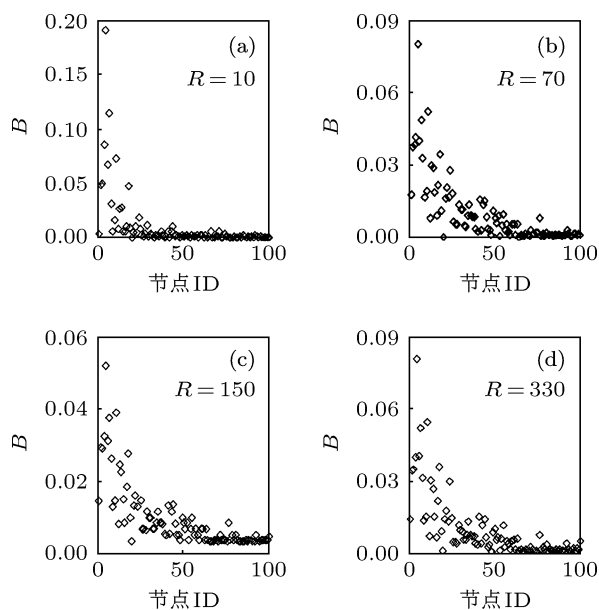


图 10 本文路由策略在不同负载量 R 下节点的介中心值分布情况 ($c = k$, $N = 100$, $m_0 = m = 4$, $\alpha = 1$, $\gamma = 2$)

4 结论

本文重点研究了网络传输过程中节点激发的

引力场形式, 建立节点的引力场方程以及传输路径对数据包的引力计算公式, 在此基础上提出了一种动态路由选择算法, 分别针对算法在不同参数取值下的路由情况进行了仿真. 仿真结果揭示了一个重要动力学现象, 即本文路由算法的性能不依赖于 α 与 γ 的取值, α 与 γ 值的变化对网络传输能力没有影响, 无论 α 与 γ 取何值 (可行域内), 算法都在较大程度地提高了网络的传输能力且网络吞吐量近似相等. 同时, 算法有效地均衡了节点的介中心值分布, 特别是节点传输能力 $c = 1$ 时, 在一定负载量下所有节点的介中心值几乎相等, 说明所有节点均得到了高效利用. 传输路径的平均长度 $\langle L_{\text{avg}} \rangle$ 随负载量 R 的增加表现出先增后减的变化趋势, 当 $c = k$ 且 $R < R_c$ 时, α , γ 值的变化对 $\langle L_{\text{avg}} \rangle$ 影响较大, 且可以通过增大 γ 值以增加路由过程对路径长度的依赖程度, 从而降低路径的平均长度.

本文研究实现了对网络传输过程中节点引力场的描述, 进一步揭示了网络传输过程的引力机理. 我们的模型能够帮助人们理解网络路由的引力作用机理, 为路由策略优化提供了一种新的研究思路. 这里我们主要以路径引力最大作为数据包传递的路由机理, 在接下来的工作中, 我们将进一步探讨在节点引力场之上的其他引力路由策略.

- [1] Newman M E J 2003 *SIAM Review* **45** 167
- [2] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D U 2006 *Physics Reports* **424** 175
- [3] Mitchell M 2006 *Artificial Intelligence* **170** 1194
- [4] Newman M E J 2010 *Networks: An Introduction* (Volume 1) (Oxford: Oxford University Press) p11–20
- [5] Wu J, Barahona M, Tan Y J, Deng H Z 2010 *Chine. Phys. Lett.* **27** 078902
- [6] Shao Z G 2010 *Appl. Phys. Lett.* **96** 073703
- [7] Toda A A 2011 *Phys. Rev. E* **83** 046122
- [8] Zager L, Verghese G 2008 *Complexity* **14** 12
- [9] Tönjes R, Masuda N, Kori H 2010 *Chaos* **20** 033108
- [10] Leyva I, Navas A, Nadal I S, Buldú J M, Almendral J A, Boccaletti S 2011 *Phys. Rev. E* **84** 065101
- [11] Daniele D M, Luca D A, Ginestra B, Matteo M 2009 *Phys. Rev. E* **79** 015101
- [12] Noh J D, Rieger H 2004 *Phys. Rev. Lett.* **92** 118701
- [13] Ramascc J J, Lama M S L, Eduardo L, Boettcher S 2010 *Phys. Rev. E* **82** 036119
- [14] Guimerà R, Díaz-Guilera A, Vega-Redondo F, Cabrales A, Arenas A 2002 *Phys. Rev. Lett.* **89** 248701
- [15] Danila B, Yu Y, Marsh J A, Bassler K E 2007 *Chaos* **17** 026102
- [16] Kawamoto H, Igarashi A 2012 *Physica A* **391** 895
- [17] Qian J H, Han D D 2009 *Acta Phys. Sin.* **58** 3028 (in Chinese) [钱江海, 韩定定 2009 物理学报 **58** 3028]
- [18] Liu G, Li Y S 2012 *Acta Phys. Sin.* **61** 108901 (in Chinese) [刘刚, 李永树 2012 物理学报 **61** 108901]
- [19] Arenas A, Díaz-Guilera A, Guimerà R 2001 *Phys. Rev. Lett.* **86** 3196
- [20] Crucitti P, Latora V, Porta S 2006 *Chaos* **16** 015113
- [21] Freeman L G 1977 *Sociometry* **40** 35
- [22] Li Q Q, Zeng Z, Yang B S, Li B J 2010 *Geomatics and Information Science of Wuhan University* **35** 37 (in Chinese) [李清泉, 曾喆, 杨必胜, 李必军 2010 武汉大学学报 (信息科学版) **35** 37]
- [23] Barabási A L, Albert R 1999 *Science* **286** 509
- [24] Echenique P, Gomez-Gardenes J, Moreno Y 2004 *Phys. Rev. E* **70** 056105
- [25] Danila B, Sun Y D, Bassler E K 2009 *Phys. Rev. E* **80** 066116

Routing strategy for complex networks based on gravitation field theory*

Liu Gang[†] Li Yong-Shu

(Faculty of Geosciences and Environmental Engineering, Southwest JiaoTong University, Chengdu 610031, China)

(Received 23 May 2012; revised manuscript received 20 June 2012)

Abstract

Using the theory of gravitational field, we study the gravitational field induced by the node in the process of the network transmission, establish the gravitational field equation, and define two parameters α and γ for adjusting the dependences of transmission data on the unblocked degree of node, the transmission capacity of node and the path length. Based on the gravitational field of node, an efficient routing strategy is proposed, and the package will be transferred along the route with maximum gravitation. In order to characterize the efficiency of the method, we introduce an order parameter η to measure the throughput of the network by the critical value of phase transition from free state to jammed state, and use the node betweenness centrality B to test the transmission efficiency of network and the congestion distribution. We simulate the network transmission efficiencies under different values of α and γ . Simulation results show that compared with the traditional shortest routing strategy, our routing strategy improves the network capacity several times, and effectively balances the distribution of the betweenness centrality of nodes, and the average path length $\langle L_{avg} \rangle$ shows a trend from ascent to descent with the increase of load amount R , and the change of the parameters α and γ nearly have no effect on the network transmission capacity, which suggests the efficiency of our routing strategy is independent of α and γ , the network capacities are approximately equal for any values of α and γ in the feasible region.

Keywords: complex network, gravitation field, routing strategy, congestion

PACS: 89.75.Hc, 89.20.Hh

* Project supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20100184110019), the 2013 Doctoral Innovation Funds of Southwest Jiaotong University and the Fundamental Research Funds for the Central Universities.

[†] E-mail: liuganggis@sina.com