动态随机最短路径算法研究*

张水舰† 刘学军 杨洋

(南京师范大学虚拟地理环境教育部重点实验室, 南京 210046)

(2011 年 12 月 30 日收到; 2012 年 4 月 27 日收到修改稿)

静态最短路径问题已经得到很好解决, 然而现实中的网络大多具有动态性和随机性. 网络弧和节点的状态及耗费不仅具有不确定性且相互关联, 弧和节点的耗费都服从一定的概率分布, 因此把最短路径问题看作是一个动态随机优化问题更具有一般性. 文中分析了网络弧和节点的动态随机特性及其相互关系, 定义了动态随机最短路径; 给出了动态随机最短路径优化数学模型, 提出了一种动态随机最短路径遗传算法; 针对网络的拓扑特性设计了高效合理的遗传算子. 实验结果表明, 文中提出的模型和算法能有效地解决动态随机最短路径问题, 可以运用到交通、通信等网络的网络流随机优化问题中.

关键词: 最短路径问题, 遗传算法, 动态随机网络

PACS: 02.10.Ox, 02.50.Cw, 05.40.−a, 07.05.TP

1 引 言

最短路径 (最优路径) 问题是图论组合优化问题中一个重要的研究内容. 在过去的几十年中, 最短路径问题的研究已经从原来的确定性最短路径问题发展到动态随机最短路径问题. 由于动态随机最短路径问题广阔的应用前景, 吸引了从通讯到交通的众多领域研究者的重视, 如在城市交通网络中, 当路段的行程时间是服从一定概率分布的随机变量时, 如何找到一条从源点 s 到目的地 d 的动态随机最短路径. 文献 [1] 还列举了动态随机最短路径在航空和通讯领域中的应用.

在实际网络中, 网络的弧和节点的耗费往往是动态和随机的 [2,3], 弧和节点的状态会随时间而改变, 具有不确定性, 在到达此弧和节点之前并不能确定其状态. 因此, 最短路径是随机事件的函数, 具有一定的概率分布. 在动态随机网络中, 各弧段和节点与邻近弧和节点是相关的, 网络上某处事件的发生往往会影响一个区域的网络弧段和节点, 而不仅仅是某条弧或某个节点. 如在交通网络中, 某处意外交通事故的发生, 由于此处交通流受阻, 必将影响邻近路段的交通状况或造成邻近路口的堵塞.

经典确定性最短路径问题已得到了深入的研究, 提出了许多高效算法 [4−6]. 经典算法可以很好地解决确定性最短路径问题, 然而对动态随机网络却不能求得最短路径, 如图 1 所示.

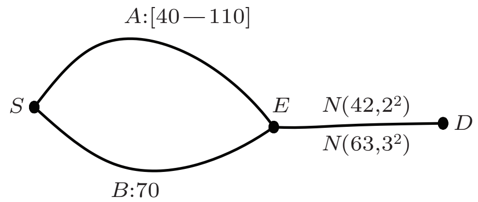

图 1 动态随机网络示例

假设在 8:00 从源节点 S 经过 E 点到达目的地 D, 想找到耗费最小的最短路径. 弧 A 的耗费服从均匀分布 [40—110], 弧 B 的耗费具有确定值 70, 节点 E 在 9:00—9:30 之间的耗费值为 7, 其他时间为 0, 弧 (E, D) 在 9:00 之前的耗费服从正态分布 $N(42, 2^2)$, 9:00 之后的耗费服从正态分布 $N(63, 3^2)$.

* 国家高技术研究发展计划 (批准号: 2007AA12Z238) 和江苏省博士后基金 (批准号: 1101016B) 资助的课题.

† E-mail: zsj_south@sohu.com

弧 A 的期望耗费值为 75, 从 S 到 E 弧 B 有较小的期望耗费值, 如果按传统最短路径算法求解, 首先找到的最优子路径是弧 B 段, 然而这条弧段对于从 S 到 D 的路径中却不是最优的. 因为经弧 A 到达 D 的期望耗费为 135, 而经弧 B 到达 D 的期望耗费为 140. 经典的最短路径算法是基于 Bellman 条件 [4], 即在一条最优路径中, 从起始点到中间节点的路径也是到达此节点的最优路径, 即每条最短路径的子路径也是最短路径, 这是经典最短路径算法的理论基础. 但动态随机网络与传统的静态网络相比, 发生了本质变化, 从上面的例子可以看出 Bellman 最优条件已不适用于弧和节点耗费具有动态性和随机性的网络.

Erdös 和 Rényi[7] 于 1960 年第一次提出了随机网络的概念. Frank[8] 对随机网络最短路径问题进行了开创性工作, 研究了在随机网络中如何确定最短路径长度的概率分布问题, 此随机网络假定路段行程时间是随机变量但不是动态的. Hall[9] 首先研究了在行程时间是随机时间依赖的交通运输网络中求解最短路径的问题, 即在网络中的弧权重是随机变量、分布函数随时间变化的情况, 证明了经典最短路径算法不能在随机时间依赖网络中找到行程最短时间期望路径. Hall 提出了一个基于动态规划的最短路径算法, 并举例说明此算法适合小规模交通网络. Fu 和 Rilett[10] 延伸了 Hall 的工作, 考虑弧耗费是时间连续随机过程的情况, 提出了行程时间期望值和方差的近似估计模型, 此模型把行程时间期望值和方差作为时间的函数来处理. Miller-Hooks 和 Mahmassani[11,12] 介绍和比较了在离散随机时间依赖最短路径问题中计算最小期望成本路径的算法, 并提出了一个新的算法.

Sigal 等 [13] 和 Kamburowski[14] 使用最短路径的概率作为优化指标, 提出了分析公式来评估这样一个指标, 其中涉及多重积分的计算且需枚举所有路径. Wellman[15] 探索了基于一阶随机优势的最优性定义, 并定义了随机优势, 提出了一种基于随机优势规则的动态规划算法来求解最优路径, 即利用随机优势删除不具有优势的路径, 扩展具有随机优势的路径, 根据效用函数决定最后输出的最短路径. 文献 [16,17] 研究了在一个完全图中只有一个结点的子集可以用来计算最短路径的情况下如何求解源节点 s 和目的节点 d 之间的最短路径.

已经有许多使用智能算法来解决路径规划问题的尝试, 如董继扬等 [18] 把神经网络运用于求单源最短路径, Alberto 等 [19] 运用蚁群算法来求解汽车的路径问题. 本文主要讨论遗传算法在最短路径问题中的应用 [20−25]. Thangiah 等 [20] 设计了一种带时间窗的最短路径遗传算法; Inagaki 等 [21] 使用了一种定长染色体的遗传算法求解最短路径, Chang 和 Ramakrishna[22] 对此进行了改进, 使用了一个可变长度染色体遗传算法, 探讨了遗传算法用于最短路径所需要的种群数量大小, 并已取得了非常好的效果; Nanayakkara 等 [23] 提出了一种大规模的路径规划遗传算法. 但上述的算法都是针对静态网络的最短路径算法, 没有考虑网络的动态性, 不适于在动态网络中求解最短路径. Davies 和 Lingras[24] 提出了动态路径遗传算法, 其中弧耗费的分布是时间的函数. Yang 等 [25] 研究了遗传算法在移动通信网络中的动态最短路径问题中的应用. 在这些动态最短路径算法中虽然考虑了网络的动态性, 但没考虑网络弧和节点的相关性. 现有最短路径遗传算法的操作算子对网络的具体特点都考虑不足, 在运算中有可能产生非法个体.

求解最短路径是一个经典的网络优化问题, 有着广泛的应用, 具有十分重要的现实意义. 传统的最短路径算法均是基于 Bellman 最优条件, 不能解决非线性路径耗费的最短路径问题, 即网络弧和节点耗费具有潜在的随机相关性、动态性问题. 从上面的分析中可看出, 虽然在确定性最短路径问题的研究上有许多成果, 但在非确定性最短路径问题上还需进一步研究.

本文旨在进一步研究动态随机网络最短路径问题, 探索如何在弧和节点耗费具有动态随机性及相互关联的网络中找到最短路径的问题, 并提出一个基于遗传算法的动态随机最短路径算法. 论文的其余部分安排如下: 在第 2 部分中, 给出了网络的表示符号以及网络最优化数学模型, 分析了网络弧和节点的相关性, 给出了概率最短路径的定义; 在第 3 部分中, 探索了解决动态随机最短路径问题的遗传算法, 设计了针对网络拓扑关系特点的遗传算子; 在第 4 部分中, 以南京市城市交通网络作为实验对象, 验证了文中提出的最短路径算法; 第 5 部分对文中的研究内容做了进一步的总结讨论, 对未来的应用进行了探讨.

2 动态随机最短路径模型

设网络由 $G = \{V, E, W\}$ 表示 (如图 2 所示), 此图由三部分组成, 其中 V 为非空节点集合, 用 i 表示节点, $V = \{1, 2, \cdots, n\}$; $E_{v\times v}$ 为弧集合 ($|E| = m$); W 是耗费 (权) 函数集合, 对图中的每个节点都赋有一个耗费值 W_i, 对图中的每条弧 $e_{ij} = (i, j) \in E$ 赋有一个耗费值 W_{ij}, 耗费值为非负. 假设图 G 不允许有循环和多重弧存在.

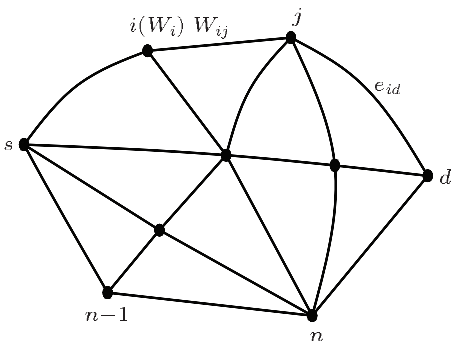

图 2 网络图示例

定义 1 路径为有限节点、弧的交替序列, 从源节点 s 到目的节点 d 的路径表示为 $r\,(s = i, e_{ij}, j, e_{jk}, k, \cdots, n = d)$, 路径不允许有圈存在.

用 $r_1\,(s, d)$, $r_2\,(s, d)$, $\cdots$, $r_k\,(s, d)$ 表示从源节点 s 到目的节点 d 的 k 条路径, 集合 $R(s, d)$ 表示从源节点 s 到目的节点 d 的所有路径的集合. 路径 $r\,(s, d)$ 的总耗费是路径所经过的各弧和各节点的耗费的总和, 以 l_r 表示, $l_r = \sum\limits_{(i,j)\in r} w_{ij} + \sum w_i + \sum w_j$, 当 W 是随机变量时, l_r 也是一随机变量; 以集合 $L\,(R_{s,d})$ 表示 $R\,(s, d)$ 中所有路径的耗费的集合, 最短路径问题就是在 $L\,(R_{s,d})$ 中找到一条耗费最少的路径. 这是个组合优化问题, 可用以下数学模型来描述:

$$l_{\min} = \min\left(\sum w_{ij}x_{ij} + \sum w_i x_{ij} + \sum w_j x_{ij}\right), \tag{1}$$

$$\left.\begin{aligned} &\sum_{i,j\in B} x_{ij} \leqslant |B|, \\ &2 \leqslant |B| \leqslant n-2, \\ &B \subset \{1, 2, \cdots, n\}, \end{aligned}\right. \tag{2}$$

$$x_{i,j} = \begin{cases} 0 & (i,j) \notin r \\ 1 & (i,j) \in r \end{cases} \quad (i, j = 1, 2, \cdots, n,\ i \neq j), \tag{3}$$

在模型中, (1) 式为目标函数, (2) 和 (3) 式为约束函数; $x_{i,j}$ 表示目标路径是否经过弧 (i, j); 目标函数 (1) 式欲使目标路径的耗费达到最小; (2) 式可以避免目标路径出现回路, 其中集合 B 是一条路径中所有节点所组成的集合的子集, $|B|$ 表示集合 B 中元素的个数; (3) 式中的决策变量 $x_{i,j} = 1$ 表示目标路径经过弧段 (i, j), $x_{i,j} = 0$ 表示目标路径不经过弧段 (i, j). 目标值 $l_{\min}$ 根据评价指标的不同可以是路径耗费的绝对值最小、或者是路径耗费的期望值最小、或者是路径耗费以最大的概率小于其他路径.

传统的最短路径问题, 假定弧和点的耗费是确定的. 然而弧和节点的状态往往是事先并不知道得这么肯定. 在网络 G 中, 从一个源节点 s, 穿过网络, 到达目的节点 d, 所经过的弧和节点的状态往往随时间发生变化, 并不是一成不变的. 弧和节点都有可能以一定概率发生堵塞, 都有两种可能的状态, 堵塞或畅通. 实际上, 网络中各弧和各节点是相关联的, 一条弧 (或节点) 堵塞可能意味着也会引起其他弧 (节点) 的堵塞. 在此种情况下, 使用概率模型会更恰当.

定义 2 设网络 $G = \{V, E, W_e(t), W_i\,(t)\}$, 其中 V 为非空的节点集合, E 为弧集合, $W_e(t)$ 为弧耗费函数集合, 弧耗费为一随机过程, $W_i\,(t)$ 为节点的耗费函数集合, 节点耗费为一时间依赖的离散随机变量, 则称 G 为动态随机网络.

在动态随机网络中, 弧耗费可建模为一个连续的随机过程, 记为 $\{W_e\,(t)\,, t \in T\}$, $W_e(t)$ 表示在时间 t 进入弧段 e 在 e 上的耗费, T 是一个连续的时间参数集. t 假定在 $t \in [t_0, t_m]$ 范围内取值, $[t_0, t_m]$ 是感兴趣的时间闭区间, t_0 是网络中结点最早的出发时间, t_m 是区间的端点, 假设 $W_e\,(t) = \infty$, $t > t_m$. 对于每一次实例 t, $W_e(t)$ 是一个连续的随机变量, 其概率密度函数表示为 $f_{w_e}\,(w_e, t)$. 当弧的流量达到一定程度时, 节点会有一个堵塞的概率 P_i, 此概率是时间依赖的, 当节点堵塞时, $W_i\,(t)$ 为一正值, 当节点畅通时, $W_i\,(t)$ 的值为 0.

当弧的流量且流速较大时, 经过节点的流量也

较大, 节点发生堵塞的概率也较大; 当其中一节点堵塞时, 各弧的流量将发生重新分配, 与此节点相连的弧将不通, 与此节点相邻的弧的耗费也将发生变化, 由此看到网络中弧和节点是相互关联的. 由网络中的弧和节点组成一个随机向量 (变量): $\boldsymbol{V} = \{W_{e_1}(t), \cdots, W_{e_m}(t), W_1(t), \cdots, W_n(t)\}$, 此向量的联合概率密度函数可表示为 $f\{w_{e_1}(t), \cdots, w_{e_m}(t), w_1(t), \cdots, w_n(t)\}$.

在弧和节点的状态具有动态性和随机性的动态随机网络中计算最短路径及其分布, 本文称之为动态随机最短路径问题. 动态随机最短路径问题与经典的最短路径问题的主要区别在于: 动态随机网络中路径的耗费值并不确定, 只知道它们的统计值及概率分布, l_r 是一随机变量, 动态随机最短路径问题就是要在从 s 到 d 的所有路径中求解最小长度 $l_{\min} = \min\{L(R_{s,d})\}$ 及其分布.

设有在某一时刻从源节点 s 到目的节点 d 的路径 $r(s = i, e_{ij}, j, e_{jk}, k, \cdots, n = d)$, 组成此路径的弧段和节点的联合概率密度函数为 $f_r = \int \cdots \int f(w_{e_1}, \cdots w_{e_m}, w_1, \cdots, w_n)\, \mathrm{d}w_{e_q} \cdots \mathrm{d}w_v$, e_q 为不在此条路径中的弧, v 为此路径不经过的节点. 假设从 s 到 d 有 u 条路径, 路径 r_i 的耗费小于其他路径的概率为 $p_i(l_i \leqslant l_1, \cdots, l_i \leqslant l_k, \cdots, l_i \leqslant l_u)$, $k \neq i$, 于是我们有以下定义.

定义 3 满足条件 $p_r = \max\limits_{i=1,2,\cdots,u} p_i(l_i \leqslant l_1, \cdots, l_i \leqslant l_k, \cdots, l_i \leqslant l_u)$, $k \neq i$ 的路径 r 称为动态随机最短路径.

此路径为概率最短路径, 对其他路径具有概率优势, 此路径期望耗费最小, 堵塞率最低. 其中

$$p_i = \int_{G(l_i \leqslant l_1, \cdots, l_i \leqslant l_u)} \cdots \int f(l_1, \cdots, l_u)\, \mathrm{d}L,$$

$f(l_1, \cdots, l_u)$ 为路径 $r_1, \cdots, r_u$ 的联合概率密度函数.

由于大规模网络的复杂性, 计算 p_r 过于复杂, 计算量巨大, 具有操作上的不可行性. 根据概率论知识我们有以下性质 [26].

性质 1 对于两随机变量 X, Y, X 的期望值 $E(X)$ 小于 Y 的期望值 $E(Y)$ 等价于 $P(X \leqslant Y) \geqslant 0.5$, 即 $E(X) \leqslant E(Y) \Leftrightarrow P(X \leqslant Y) \geqslant 0.5$.

证明 假设当 $E(X) \leqslant E(Y)$ 时, 有 $P(X \leqslant Y) \leqslant 0.5$, 然而这是不可能的, 如假设有两独立的随机变量 X 和 Y, 它们的概率密度分别为

$$f_X(x) = \begin{cases} 1/2, & 6 < x < 8 \\ 0, & \text{其他} \end{cases},$$

$$f_Y(y) = \begin{cases} 1/8, & 4 < y < 12 \\ 0, & \text{其他} \end{cases},$$

则 (X, Y) 的联合概率密度为

$$\begin{aligned} f(x, y) =& f_X(x) f_Y(y) \\ &= \begin{cases} 1/16, & 6 < x < 8,\ 4 < y < 12 \\ 0, & \text{其他} \end{cases}, \end{aligned}$$

$E(X) = 7 < E(Y) = 8$, 而 $P(X \leqslant Y) = \iint_{\Omega} f(x, y)\, \mathrm{d}x\mathrm{d}y = 0.625 > 0.5$ (其中 Ω 是 $6 < x < 8$, $4 < y < 12$, $x \leqslant y$ 所围成的区域), 与 $P(X \leqslant Y) \leqslant 0.5$ 矛盾. 因此有 $E(X) \leqslant E(Y) \Rightarrow P(X \leqslant Y) \geqslant 0.5$, 同理有 $P(X \leqslant Y) \geqslant 0.5 \Rightarrow E(X) \leqslant E(Y)$, 所以 $E(X) \leqslant E(Y) \Leftrightarrow P(X \leqslant Y) \geqslant 0.5$.

此性质表示当路径 X 的耗费期望值 $E(X)$ 小于路径 Y 的耗费期望值 $E(Y)$ 时, 路径 X 对路径 Y 具有概率优势, 因此我们可以通过比较各路径的期望值来求得具有最大概率优势的路径. 在大规模网络中, 要找到概率优势路径是困难的, 需要巨大的计算量. 从引言部分可知, 经典的最短路径算法不能在动态随机网络中找到最短路径, 动态随机最短路径问题是一个 NP-hard 问题 (对于问题的一个解, 不能在多项式时间内验证这个解是否为正确的问题称为 NP-hard 问题), 在某种意义上没有多项式时间算法, 需要提出新的有效算法来求解最优路径. 遗传算法适合大空间的搜索, 能在效率与精度之间取得良好的折中, 已有研究人员应用遗传算法于最短路径问题中, 取得了很好的效果, 显示出遗传算法在此类问题中的优越性. 下面我们提出一种最短路径遗传算法以求解动态随机最短路径问题.

3 最短路径遗传算法

遗传算法具有解决复杂问题的能力, 已成功地应用在各种实际应用中 [22]. 遗传算法 (genetic algorithm) 是美国 Michigan 大学的 Holand[27] 于 1975

年首先提出. 遗传算法将计算机科学与遗传学原理相结合, 根据优胜劣汰的原则, 使所要解决的问题从开始解逐渐逼近最优解或准最优解. 作为一种全新的随机优化与搜索方法, 遗传算法具有良好的并行性、全局优化性、稳健性、很强的通用性、良好的可操作性和简单性等显著特点, 对很多优化问题能够较容易地给出令人满意的解, 已在科学研究和工程最优化领域中展现出独特魅力, 在过去几十年中遗传算法得到了良好的发展. 遗传算法的运算过程如图 3 所示.

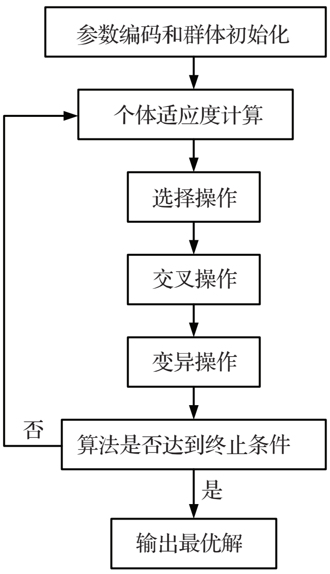

图 3 遗传算法基本流程图

我们解决动态随机最短路径问题应该建立一种高效搜索技术以减少搜索到最短路径的时间耗费, 提高效率, 而不是遵循列举所有可能路径的思路. 本文提出的最短路径算法正是基于遗传算法优越的性能, 查找潜在的最短路径, 而不是枚举所有路径.

3.1 染色体编码规则

染色体编码有二进制编码和实数编码两种形式. 在网络中节点的排列顺序正好组成一条路径, 采取有序的实数编码方式更为合理, 染色体由表示路径节点的正整数序列组成, 染色体的基因为路网节点, 每个基因位置代表一个节点在路径中的顺序.

染色体第一个基因是源节点, 最后一个基因位置是目的节点. 染色体的长度在 2 和 N 之间, N 是网络中的节点总数. 由于不允许出现圈, 因此一条路径的节点数不会多于 N. 染色体的长度是可变的, 但它不应该超过最大长度, 这里是网络中的节点总数. 由于各条路径所要经过的节点都不一样, 所以采用不等长可变染色体的实数编码形式, 这种编码形式很适合相同起点和终点之间的不同连通路径具有不同节点数的情况. 如采用定长染色体编码, 将产生巨大的搜索空间, 算法的效率将是无法保证的. 图 4 中从节点 1 到节点 5 的有一路径 1→2→7→10→4→5 (本文以有向网络为例, 算法可以很容易移植到无向网络中), 图 5 为其染色体编码.

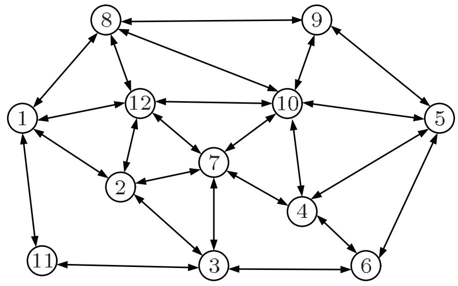

图 4 网络示意图 (有向网络)

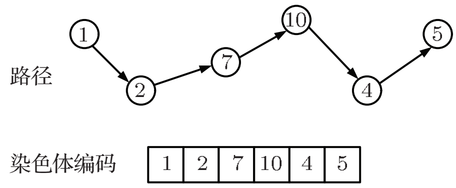

图 5 染色体编码示例

3.2 种群的初始化

遗传算法具有很强的鲁棒性, 这是建立在初始种群多样性的基础之上的. 在遗传算法中初始种群的好坏将直接影响算法的收敛速度和收敛结果.

由于跟一个节点相连的弧段通常有多条 (节点出度大于 2), 对于节点成千上万的大规模网络, 找到一条从源节点到目的节点的可能路径是非常困难的. 如果从源节点开始, 随机选取一个邻接节点作为下一个节点, 持续此操作, 直到目的节点, 这种方法需要花费很长的时间来找到一条路径, 效率太低, 基本上不可行. 为此, 我们设计了下面的启发式方法来进行种群的初始化, 此方法具有更高的效率, 且可以避免回路和死点, 其步骤简要描述如下:

步骤 1 根据先验知识, 给每一弧段赋一个先

验权值;

步骤 2　以起点作为染色体的第一个基因, 在与以此节点为尾节点的弧段中, 排除头节点的出度为 1 的弧段, 然后选取权值最小的弧段的头节点且作为此路径的下一个节点 (第二个基因), 并标记该节点;

步骤 3　以第二个节点为起点按步骤 2 的操作求第三个结点 (没标记的节点) 作为第三个基因, 按此方法一直到目的节点, 即求得一个初始染色体;

步骤 4　如果种群数量已到达所需要的数量, 则终止; 否则重复步骤 2 和步骤 3 求下一个染色体.

3.3　选择算子

选择操作可以让高品质染色体复制到下一代有更多的机会, 以提高整个种群质量 [28]. 选择操作使得算法集中在有希望的解空间中进行搜索. 通常的选择机制有轮盘赌选择、$(\mu+\lambda)$ 选择、竞争选择等.

适应值函数被用来评价潜在的最优解. 此处适应值低的解是好的解. 定义适应值函数为从源节点 s 到目的节点 d 所经过的弧段和节点的耗费期望值之和, 第 k 个染色体的适应值函数可表示为

$$f(k)=\sum E\left(W_i\left(t\right)\right)+\sum E\left(W_e\left(t\right)\right). \tag{4}$$

在此采用 Holland 提出的轮盘赌选择方式进行个体的选择, 其基本原理是根据每个染色体适应值的比例来确定该个体的选择概率.

3.4　交叉算子

交叉操作是遗传算法中另一个非常重要的部分, 它主要影响算法的搜索效率. 路径交叉操作是交换两个染色体中的子路径. 种群初始化操作使得染色体都拥有相同的源节点和目的节点. 被选择到进行交叉操作的染色体会有两种情况, 一是两个父染色体有相同的节点 (除源节点和目的节点外); 二是两个父染色体没有相同的节点.

对第一种情况采取在相同节点进行交叉操作, 如有多个相同节点, 则在这几个节点中随机选择一个节点作为交叉位置. 以图 4 所示网络为例, 此类交叉操作如图 6(a) 所示.

由于现实网络往往是不完全网络, 对第二种情况, 由于有可能在交叉操作中产生非法的个体, 故要对非法个体进行修正. 以图 4 示意的网络为例, 假设以节点 1 为源节点和节点 5 为目的节点, 选择到两条有效路径 1→8→10→9→5 和 1→2→3→7→4→6→5 作为双亲染色体. 如图 6(b) 所示, 随机选择一基因位置{10, 3}进行交叉操作, 交叉操作后产生的子代 2 是不合法的路径, 节点 3 和节点 9 之间没有弧段相连接, 需进行修正. 为此按照 5.2 中介绍的启发式方式找到一条从节点 3 到节点 9 的有效路径, 使子代 2 染色体为合法路径.

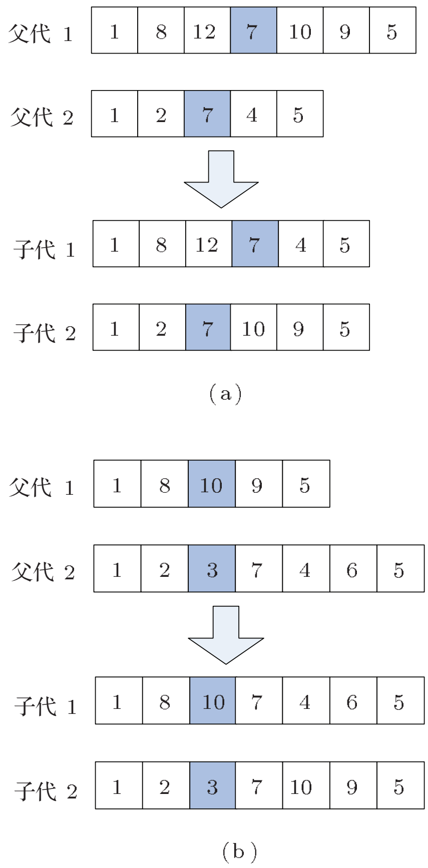

图 6　交叉操作示意图

3.5　变异算子

路径变异操作运算是从一染色体产生另一染色体. 考虑到不完全网络拓扑关系的约束, 变异算子的运算会产生非法个体, 故要对变异算子进行改进. 在此采取以下的变异方式, 分两步进行:

1) 以一定概率随机选择变异的基因位置;

2) 此基因连接前后两基因组成一条子路径, 用另一条子路径替换此子路径.

为了说明以上步骤, 仍以图 4 示意的

网络为例, 假设上述交叉操作所产生的子代 1→8→10→7→4→6→5 中的节点 4 被选择作为变异的基因, 去掉节点 4, 然后再在节点 7 和节点 6 之间找到一条满足约束条件的合法路径, 如图 7 所示.

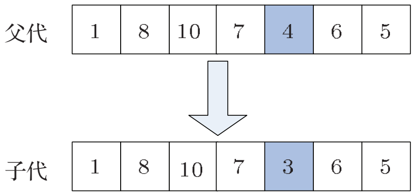

图 7　变异操作示意图

3.6　算法步骤

以下是最短路径遗传算法的形式化描述:

1) 确定节点阻抗函数 $W_i\,(t)$ 和网络弧阻抗函数 $W_e\,(t)$;

2) 种群的初始化;

3) 计算种群各染色体的适应值 $eval(k)$;

4) 进行选择操作;

5) 进行交叉操作;

6) 进行变异操作;

7) 满足算法终止条件, 算法终止, 否则转 3).

4　实　验

考虑一个由有限点集和弧集组成的有向图表示的交通网络, 此交通网络中的每条弧和每一个出度不为零的节点都有一个与之关联的广义成本, 这可能是行程时间或成本与距离的综合或是由于堵塞引起的延误等. 不失一般性, 本文将使用路段的行程时间和节点的延误来表征这个广义成本. 假设此交通网络上路段的行程时间是一随机过程, 路段行程时间的概率分布依赖于何时到达此路段起点的时间; 出度不为零的节点的延误也服从一定的概率分布, 此概率分布依赖于何时到达该节点. 节点的延误服从离散分布, 路段行程时间服从正态分布 [29]. 如在时间 7:30—8:30 之间, 北京东路与太平北路相交的路口的堵塞率为 0.1, 引起时间延误为 10 min, 其期望延误为 1 min; 路段九华宾馆 — 白马公园的行程时间服从正态分布 $N(5, 3^2)$.

实验在南京市的交通网络上进行, 此网络由 2668 条路段和 1677 个节点组成 (如图 8 所示). 上午高峰期 (上午 6:30— 上午 8:30) 被选定为研究时间区域. 由于缺乏实时的交通状况统计数据, 只能创建一个简化的动态随机网络, 分时段表示各路段的行程时间以及节点的延误, 以每 12 min 为时间间隔表示动态随机行程时间模式, 研究时间区域共分为 10 个时段.

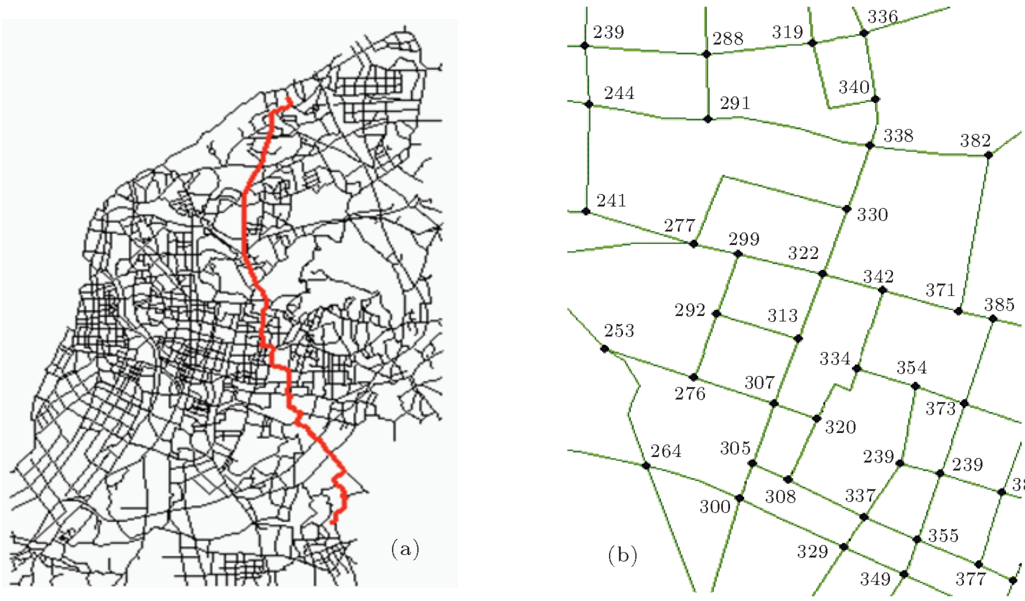

图 8　南京市交通网络图　(a) 最短路径示例 (粗线为南京电瓷厂到下章村的最短路径); (b) 交通网络局部放大图 (图中数字为节点编号)

算法是在 GIS 软件平台 ArcGIS10 的支持下, 在 VS. NET 开发环境中用开发工具 C# 编程仿真实现. 软件平台 ArcGIS10 具有良好的矢量地图处理能力, 可以方便地生成一个交通网络图 (见图 8). 在此交通网络中根据路段的距离及交通状况等因素给每条路段的不同方向不同时段赋以不同的耗费分布函数, 以求取期望耗费值 $W_e(t)$; 给每个节点在不同时段赋以一个堵塞概率 $p\,(p \geqslant 0)$, 以及一个延迟值, 以求取节点期望耗费值 $W_i\,(t)$. 利用 VS. NET 开发环境中的 ArrayList 类存储染色体编码, 以进行各项遗传操作. 选取遗传算法参数为: 种群数量 $pop\_\mathrm{size} = 50$, 交叉概率 $p_\mathrm{c} = 0.85$, 变异概率 $p_\mathrm{m} = 0.02$, 最大进化代数 $\mathrm{Maxgen} = 200$. 在交通网络中随机选取了 9 对源 - 目的节点, 每对源 - 目的节点选取不同的初始种群计算了 3 次, 试验结果如表 1 所示. 试验计算结果都能收敛, 说明此算法有良好的收敛性. 从表 1 的试验结果中可以看到算法计算到最优路径的平均计算时间少于 3 s, 说明算法有很高的效率.

表 1 试验结果

序号	源 - 目的节点对		收敛值/min	平均计算时间/s
	源节点	目的节点		
1	南师随园	仙鹤山庄	35	2.5
2	南京电瓷厂	下章村	51	3.6
3	审计学院	青马村	37	2.1
4	兴隆村	西场坊	42	3.6
5	南京火车站	卡子门广场	30	2.2
6	南京西站	理工大学	36	2.8
7	朝天宫	尧化新村	34	1.9
8	林业大学	典雅居	25	1.3
9	张家湾	乌龙山公园	68	4.2

5 结 论

本文研究了动态随机最短路径问题, 分析了网络的动态性和随机性, 把网络弧和节点耗费建模为随机过程, 分析了网络弧和节点的相关性, 在此基础上提出了概率最短路径的定义. 探讨了遗传算法应用于最短路径问题的基本方法和策略, 提出了一种最短路径遗传算法以解决动态随机最短路径问题. 得出以下几个结论.

1) 传统确定性算法没有考虑网络的动态性或随机性, 现实网络大多具有动态性和随机性, 网络弧段和节点的耗费是动态且是随机的, 如交通网络、交通状况随时都发生改变. 经典最短路径算法假定路径的耗费是固定的, Bellman 提出的路径最优化原理可以用来解决确定性最短路径问题, 然而此原理不能很好地解决动态随机最短路径问题, 经典算法不能在动态随机网络中找到最短路径.

2) 网络中弧和节点的状态往往是相互关联的, 以往最短路径算法多没有考虑弧和节点的相关性 [16], 与传统网络模型相比, 考虑弧和节点相关性的概率最短路径更具有普遍意义. 由于大规模网络的复杂性, 直接求解概率最短路计算量巨大, 具有不可行性, 根据期望值与概率最短路径的关系, 我们可以通过比较期望值来找到具有概率优势的概率最短路径.

3) 遗传算法具有大空间搜索能力, 很适合用来求解大规模网络的组合优化问题, 已往最短路径遗传算法对网络的具体特征考虑不足 [24,25], 需根据网络的拓扑关系特点改进遗传算子. 实验证明本文所提出的算法具有良好的收敛性和运行效果, 能有效地解决动态随机最短路径问题, 说明遗传算法是解决此类问题的有效工具.

4) 本文提出的动态随机最短路径模型和算法可以应用到通信网络、交通网络等领域的网络流优化中, 如可以用在智能交通系统中对车辆进行路径的动态诱导, 使出行者能快速便捷地到达目的地, 以减少车辆在交通网络上的停留时间, 对合理引导交通流及缓解拥挤的交通具有一定的积极意义.

此外应该指出, 本文的研究是基于假设网络弧段及节点耗费是可预测的, 其具有一定的概率分布形式, 对于不可预测的情况下求解最短路径将更困难, 这些问题需要进一步深入研究以得出更一般性的结论.

[1] Peer S K, Dinesh K S 2007 *Comput. Math. Appl.* **53** 729

[2] Li S B, Wu J J, Gao Z Y, Lin Y, Fu B B 2011 *Acta Phys. Sin.* **60** 050701 (in Chinese) [李树彬, 吴建军, 高自友, 林勇, 傅白白 2011 物理学报 **60** 050701]

[3] Wang K, Zhou S Y, Zhang Y F, Pei W J, Liu Q 2011 *Acta Phys. Sin.* **60** 118903 (in Chinese) [王开, 周思源, 张毅锋, 裴文江, 刘茜 2011 物理学报 **60** 118903]

[4] Bellman E 1958 *Quart. Appl. Math.* **16** 87

[5] Dijkstra E W 1959 *Numer. Math.* **1** 269

[6] Dreyfus S 1969 *Operat. Res.* **17** 395

[7] Erdös P, Rényi A 1960 *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17

[8] Frank H 1969 *Operat. Res.* **17** 583

[9] Hall R W 1986 *Transport. Sci.* **20** 182

[10] Fu L, Rilett L R 1998 *Transport. Res.* B **32** 499

[11] Miller-Hooks E D, Mahmassani H S 1998 *Comput. Operat. Res* **25** 1107

[12] Miller-Hooks E D, Mahmassani H S 2000 *Transport. Sci.* **34** 198

[13] Sigal C E, Pritsker A A B, Solberg J J 1980 *Operat. Res.* **28** 1122

[14] Kamburowski J 1985 *Operat. Res.* **22** 696

[15] Wellman M P 1995 *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence* Montreal, Quebec, Canada, August 18–20, 1995 p18

[16] Jaillet P 1992 *Networks* **22** 589

[17] Fan Y Y, Kalaba R E, Moore J E 2005 *Comput. Math. Appl.* **49** 1549

[18] Dong J Y, Zhang J Y, Chen Z 2007 *Acta Phys. Sin*. **56** 5013 (in Chinese) [董继扬, 张军英, 陈忠 2007 物理学报 **56** 5013]

[19] Alberto D V, Roberto M, Norman C, Andrea E R, Gambardella L M 2008 *Eur. J. Operat. Res.* **185** 1174

[20] Thangiah S R, Nygard K, Juell P 1991 *Proceedings of the 7th IEEE Conference on Artificial Intelligence Applications* Miami, USA, February 24–28, 1991 p422

[21] Inagaki J, Haseyama M, Kitajima H 1999 *Proc. IEEE Int. Symp. Circuits and Systems* Orlando, USA, May 30–June 2, 1999 p137

[22] Chang W A, Ramakrishna R S 2002 *IEEE Trans. Evol. Comput.* **6** 566

[23] Nanayakkara S, Srinivasan D, Lup L, Xavier G, Elizabeth T, Ong S H 2007 *IEEE Congress on Evolutionary Computation* Singapore, September 25–28, 2007 p4469

[24] Davies C, Lingras P 2003 *Eur. J. Operat. Res.* **144** 27

[25] Yang S, Cheng H, Wang F 2010 *IEEE Trans. Syst. Man Cybernet.* C **40** 52

[26] Charles J S 1996 *A Course in Probability and Statistics* (California: Duxbury Press) p81

[27] Holland J H 1975 *Adaptation in Natural and Artificial Systems* (Michigan: University of Michigan Press) p22

[28] Thomas B W, White C C 2007 *Eur. J. Operat. Res.* **176** 836

[29] Bell M G H, Iida Y 1997 *Transportation Network Analysis* (Chichester: John Willey and Sons) p17

Dynamic stochastic shortest path algorithm*

Zhang Shui-Jian† Liu Xue-Jun Yang Yang

(*Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing* 210046, *China*)

(Received 30 December 2011; revised manuscript received 27 April 2012)

Abstract

The static shortest path problem has been solved well. However, in reality, more networks are dynamic and stochastic. The states and costs of network arcs and nodes are not only uncertain but also correlated with each other, and the costs of the arcs and nodes are subject to a certain probability distribution. Therefore, it is more general to model the shortest path problem as a dynamic and stochastic optimization problem. In this paper, the dynamic and stochastic characteristics of network nodes and arcs and the correlation between the nodes and arcs are analyzed. The dynamic stochastic shortest path is determined. The dynamic stochastic optimization model of shortest path is provided, and a shortest path genetic algorithm is proposed to solve dynamic and stochastic shortest path problem. The effective and reasonable genetic operators are designed according to the topological characteristics of the network. The experimental results show that this algorithm can be used to effectively solve the dynamic stochastic shortest path problem. The proposed model and algorithm can be applied to the network flow optimization problem in transportation , communication networks, etc.

Keywords: shortest path problem, genetic algorithm, dynamic and stochastic network

PACS: 02.10.Ox, 02.50.Cw, 05.40.−a, 07.05.TP

* Project supported by the National High Technology Research and Development Program of China (Grant No. 2007AA12Z238) and the Postdoctoral Science Foundation of Jiangsu Province, China (Grant No. 1101016B).

† E-mail: zsj_south@sohu.com