

文章编号: 1007-6735(2011)03-0274-05

单向限量最速网络消息传播模型及其进化算法

何胜学

(上海理工大学 管理学院, 上海 200093)

摘要: 提出了单向限量式最速网络消息传播问题,建立了该问题的数学模型,并给出了相应的模拟进化求解算法.通过分析单向限量式最速网络消息传播问题的特征,包括决策变量的特点、决策的网络时空影响特殊模式及网络消息分布状态特点,构建了问题的最优化模型.利用决策变量的二元取值特点和单一轮次信息交互模式的相对独立性,设计了操作灵活的遗传算法的复制、交叉和变异算子,实现了模型的模拟进化求解.数值算例验证了模型和算法的有效性.最后总结了最速网络消息传播问题的主要可扩展研究方向.

关键词: 系统工程; 网络优化; 消息传播; 遗传算法

中图分类号: O 221; O 157.6 **文献标志码:** A

Fastest network message one-way spreading model with bounds of transitive messages and its evolutionary algorithm

HE Sheng-xue

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The fastest network message one-way spreading problem with bounds of transitive messages was introduced. The problem was formulated with strict mathematics. The corresponding simulated evolutionary algorithm was provided. Through analysing the characteristics of the problem, including the features of decision variables, the special pattern of the spatial and temporal impacts of decision-makings and the distribution features of network messages, the optimization model of the problem was built. Taking advantage of the binary feature of decision variables and the relative independence of the pattern of single round information interchanging, the reproduction operator, the crossover operator and the mutation operator of genetic algorithm were designed, that can be manipulated flexibly. So the simulated evolutionary solution of the model was achieved. The numerical example demonstrates the effectiveness of the model and the algorithm. The main extensible research directions with regard to the problem were summed up.

Key words: systems engineering; network optimization; message spreading; genetic algorithm

收稿日期: 2011-03-23

基金项目: 上海市优秀青年教师基金资助项目(slg08018); 上海市教育委员会科技创新项目(10YS105)

作者简介: 何胜学(1976-),男,讲师.研究方向: 交通网络建模. E-mail: lovellhe@126.com.

随着网络信息科技的发展,网络交流已成为人们日常生活的一部分.如何快速地将各种信息通过网络进行传播成为网络通讯、系统管理优化以及军事信息技术领域的研究热点问题.及时准确的大规模信息互通已成为解决诸如战场单兵协同作战、局域网络最优布线以及易损网络重建等问题的瓶颈^[1-4].但是,相关理论研究的滞后制约了这些领域的进一步发展.

通过对单向限量最速网络消息传播问题加以严格的数学建模,分析该基本问题的基本约束特征,以及设计相应的模拟进化求解算法,为最速网络消息传播问题的深入研究建立相应的理论基础.

1859年达尔文创立进化论,成为人类文明历史上的一个里程碑.遵循“生存竞争,优胜劣汰”的进化原则而设计的进化算法主要包括遗传算法、遗传规划、进化策略和进化规划这4种典型方法^[5-6].由于进化算法的自适应式搜索寻优特征、结构化并行的计算特征、全局优化性及广泛的适用性,目前已经被成功地应用于各种工程技术领域^[5-11].鉴于单向限量式最速网络消息问题的变量离散化特征、多阶段非线性优化特征及复杂的网络拓扑结构,模拟进化算法中的遗传算法成为解决该问题的首选工具.进化算法中复制、交叉和变异操作能较好地适应网络消息传播的交互特征,从而使设计的具体算法不仅构造简洁,而且实施方便.

1 最速网络消息传播问题

假设存在 n 个信息源,每个信息源在信息交互未发生前,各有一个消息须向其他的 $n-1$ 个信息源传达.任意一对信息源之间均存在一条信息交互通道.一个信息源至多只能同时与一个其他信息源通过信息交互通道传达消息,信息交互是单向的,且一次交互中有消息通讯量的上限限制.假设每一次交互的时长相等,至少需要进行多少次交互才能使所有消息最快传至所有信息源,这就是单向限量最速网络消息传播的基本问题.

现给出主要的参变量定义:

i, j 表示信息源 i 和 $j, i, j \in \{1, 2, \dots, n\}, n$ 为信息源总数; t 表示是第 t 轮信息交互, $t \in \{1, 2, \dots, 2n\}$; x_{ij}^t 表示信息源 i 和 j 在第 t 轮信息交互时是否建立交互的指标,存在信息源 i 向信息源 j 传播消息的交互时,其值为 1, 否则,为 0; I 表示所有的消息构成的集合,即 $\{I_1, I_2, \dots, I_n\}$, 这里 I_j 表示

初始阶段信息源 j 打算向外传递的信息; U_i^t 表示第 t 轮信息交互结束后信息源 i 已知的消息集合, $U_i^0 = \{I_i\}$ 表示信息源 i 起始状态的消息集合; δ^t 表示第 t 轮信息交互的费用系数,有 $0 < \delta^t < \delta^{t+1}$, $\forall t \in \{1, 2, \dots, 2n-1\}$ 成立.

如果 n 个信息源间每次交互只传递自己初始态时想传递的消息,那么双向的最速网络消息传播问题退化为类似 n 个足球队间的锦标赛问题.易知,此时至少需进行 $n(n-1)/2$ 次交互.而每一轮交互中, n 个信息源间至多可进行 $\lfloor n/2 \rfloor$ 次交互. n 为偶数时, $\lfloor n/2 \rfloor$ 等于 $n/2$; 否则, $\lfloor n/2 \rfloor = (n-1)/2$. 考虑到 $n \lfloor n/2 \rfloor \geq n(n-1)/2$, 因此,可设单向限量最速网络消息传播时, t 最大取值为 $2n$.

最速网络消息传播基本问题的最优化模型为

$$\min Z = \sum_{t=1}^{2n} (\delta^t \sum_{i=1}^n \sum_{j=1}^n x_{ij}^t) \quad (1)$$

$$\text{s. t. } x_{ii}^t = 0, \forall i \in \{1, 2, \dots, n\}, \quad \forall t \in \{1, 2, \dots, 2n\} \quad (2)$$

$$x_{ij}^t \neq x_{ji}^t, \quad \forall i, j \in \{1, 2, \dots, n\}, \quad \forall t \in \{1, 2, \dots, 2n\} \quad (3)$$

$$\sum_{i=1}^n x_{ij}^t \leq 1, \quad \forall j \in \{1, 2, \dots, n\}, \quad \forall t \in \{1, 2, \dots, 2n\} \quad (4)$$

$$\sum_{j=1}^n x_{ij}^t \leq 1, \quad \forall i \in \{1, 2, \dots, n\}, \quad \forall t \in \{1, 2, \dots, 2n\} \quad (5)$$

$$U_j^t = U_j^{t-1} \cup R(q, U_i^{t-1} \cap \bar{U}_j^{t-1}), \text{ 如果}$$

$$x_{ij}^t = 1, \forall i, j, \forall t \in \{1, 2, 3, \dots, 2n\} \quad (6)$$

$$U_i^{t-1} \subseteq U_i^t, \quad \forall i \in \{1, 2, \dots, n\}, \quad \forall t \in \{1, 2, \dots, 2n\} \quad (7)$$

$$U_i^0 = \{I_i\}, \quad \forall i \in \{1, 2, \dots, n\} \quad (8)$$

$$U_i^{2n} = I, \quad \forall i \in \{1, 2, \dots, n\} \quad (9)$$

$$x_{ij}^t \in \{0, 1\}, \quad \forall i, j, \forall t \in \{1, 2, 3, \dots, 2n\} \quad (10)$$

式中, Z 为最小化加权后各轮信息交互的总次数; q 为单次交互中可以传递的消息的最大数目, 满足 $1 \leq q \leq n$; $\bar{U}_j^{t-1}$ 表示信息源 j 在 $t-1$ 轮信息交互后, 还未了解的信息集合; 交集 $U_i^{t-1} \cap \bar{U}_j^{t-1}$ 代表信息源 j 在 $t-1$ 轮信息交互后, 还未了解的信息中, 已经被信息源 i 了解的信息构成的集合; 而集合 $R(q, U)$ 表示如果集合 U 的元素个数大于 q , 从集合 U 随机取出 q 个元素构成的新集合, 而如果集合 U 的元素个数小于或等于 q , $R(q, U)$ 等于集合 U .

由目标函数式(1)和约束式(2)~(10)共同构成单向限量最速网络消息传播基本问题的规划模型. 由于 $0 < \delta^t < \delta^{t+1}$, $\forall t$ 成立, 因此, 信息源间的有效交互发生得越早越好, 从而体现最速最简洁的信息交互要求.

约束式(2)表示每轮的信息交互过程中单个信息源不存在与自身的信息交互. 约束式(3)表示两个信息源之间如发生信息交互, 则交互是单向的. 约束式(4)和式(5)共同限制了每一个信息源在任何一轮的信息交互过程中至多只能进行一次交互. 约束式(7)表示随着信息交互的逐轮进行, 任意信息源所知的消息只能越来越多. 约束式(8)和式(9)分别表示信息源掌握消息的初始状态和最终状态. 约束式(10)表明任意可能的交互要么发生, 要么没有发生. 约束式(6)表示如果某轮信息交互中两个信息源间发生信息交互, 则该轮信息交互使得两个信息源间形成单向限量式的已知消息交互.

在上面给出的优化模型基础上, 扩展建立其他相关的最速网络消息传播问题的优化模型将非常方便. 例如考虑信息交互的双向性, 相应的优化模型只需将约束式(3)的不等号变为等号即可. 当考虑多样化的信息源初始状态和最终状态时, 只需对约束式(8)和式(9)进行修改即可. 而如果每次信息交互的最大消息数无上限时, 可通过修改约束式(6)中的 q 值实现. 综上, 由式(1)~(10)构成的单向限量最速网络消息传播问题的优化模型是最速网络消息传播问题族的建模基础.

2 模拟进化解法设计

应用模拟进化计算中的遗传算法需解决的关键问题包括: a. 将问题的解进行编码, 构成遗传代次中的个体; b. 生成合理的初始群体; c. 设计方法用以计算遗传个体的适应度; d. 设计相应的复制、交叉和变异算子. 现对上述问题进行具体分析, 然后给出完整的模拟进化求解算法实施步骤.

最速网络消息传播问题的变量 x_{ij}^t 只能取值 0 或 1, 正好符合遗传算法中染色体的二进制编码要求. 同时由约束式(2)和式(3)可知, 任何一轮的信息交互只需由满足条件的变量 x_{ij}^t , $\forall i < j$ 构成的 $n(n-1)/2$ 个变量组合就可代表该轮的信息交互模式. 因此, 将所有 $2n$ 轮的上述变量组合依次合并, 就可以构成一个满足算法要求的个体编码.

生成一个满足约束式(2)~(5)的可行消息交互

模式, 即遗传个体, 需要 4 个步骤. 首先, 生成 n 个信息源的一个随机排列. 可采取遗传算法的轮盘选择原理逐次减少信息源构成等分轮盘进行信息源选择, 从而形成 n 个信息源的一个随机排列. 接着, 从前向后依次从生成的随机排列中选择两个信息源, 构成一个可能的有向交互. 总共有 $[n/2]^-$ 对可能的信息源对, 即交互. 随后, 对每一个可能交互生成一个在 0 和 1 之间服从均匀分布的随机变量值 $\epsilon_{[0,1]}$. 比较 $\epsilon_{[0,1]}$ 与依据当前的轮次 t 给定一个判别值 ϵ^t 的大小, 如果 $\epsilon_{[0,1]} > \epsilon^t$, 使对应的可能交互实现; 否则, 对应的可能交互不予实现. ϵ^t 的值大于 0, 小于 1. ϵ^t 随着轮次 t 的增大而增大. 最后, 将所有轮次的交互模式组合, 即得到一个满足约束式(2)~(5)的可行消息交互模式. 重复上述步骤, 可以得到遗传算法的初始群体.

遗传个体的适应度计算分为两步. 第 1 步依据问题的目标函数式(1)计算总的交互次数. 由于不是所有的个体(即信息交互模式)均能使得系统的终态满足式(9), 因此需要在第 2 步计算一个相应的惩罚项. 对惩罚项的计算分两种情况. 第 1 种情况为至少存在一个信息源掌握所有消息. 因为, 未能掌握所有消息的信息源可以通过与已经掌握所有消息的信息源进行交互, 从而了解所有消息, 所以未能掌握所有消息的信息源的未了解的消息总数可以作为惩罚项, 加权后加入该遗传个体的适应度. 第 2 种情况为所有信息源在 n 轮交互结束后, 均未能了解所有的消息. 此时, 首先确定生成一个掌握所有消息的信息源需进行的一个可行的交互次数, 即找到掌握消息最多的信息源, 用消息总数 n 减去该信息源已经掌握的消息数, 并除以消息传播上限 q . 将这个可行的交互次数与 $(n-1)$ 个信息源未知消息总量除以消息传播上限 q 的值相加, 得到的结果可以作为惩罚项, 加权后加入该遗传个体的适应度.

针对最速网络消息传播问题设计的遗传算法的复制算子与常用的复制算子无异. 既可以采取简单的依据适应度大小的比例删除法, 也可以利用轮盘选择原理进行选择. 但与常用的交叉和变异算子不同, 对于最速网络消息传播问题, 为了满足生成的新个体的可行性, 交叉和变异应以交互的轮次为单位, 进行对应轮次的交互模式互换和以轮次为单位的染色体变异. 这样可以使新生成的个体满足除约束式(9)外的其它基本约束.

求解最速网络消息传播基本问题的遗传算法基本步骤如下:

步骤 1 初始化. 确定种群规模 N , 选择复制比例 θ , 交叉概率 P_c , 变异概率 P_m 和最大迭代次数 $N = 100$; 随机生成 $\theta = 0.85$ 个个体作为初始种群 $P_c = 0.75$; 置迭代指标 $P_m = 0.15$ 为 0.

步骤 2 个体适应度计算. 考虑不满足约束式 (8) 的加权惩罚项, 计算常数 $T = 50$ 中各个体的适应度.

步骤 3 种群进化.

步骤 3.1 (复制) 依据适应度的大小选取占比例为 θ 的个体组成新的种群 $G(\tau)$;

步骤 3.2 (交叉) 从 $G(\tau)$ 中选择出 $M/2$ 对母体 ($M \leq N$), 依据概率 P_c 执行交叉, 形成 M 个中间个体;

步骤 3.3 (变异) 对 M 个中间个体分别独立依概率 P_m 执行变异, 形成 M 个候选个体;

步骤 3.4 (选择) 计算 M 个候选个体的适应度后, 将其加入种群 $G(\tau)$; 依据适应度大小从种群 $G(\tau)$ 中选择 N 个个体作为新种群 $X(\tau + 1)$.

步骤 4 终止检验. 如果 $\tau + 1 \leq T$, 则输出 $X(\tau + 1)$ 中具有最大适应度的个体作为最优解; 否则, 令 $\tau := \tau + 1$, 转步骤 2.

上述求解步骤中终止准则也可以采取其他的形式, 如当连续多次迭代后, 目标函数得不到改进, 即终止计算.

3 算例分析

分析只有 8 个信息源的单向限量式最速网络消息传播问题. 设定种群规模 $N = 100$, 复制比例 $\theta = 0.85$, 交叉概率 $P_c = 0.75$, 变异概率 $P_m = 0.15$, 最大遗传代次 $T = 100$, 单向的消息最大交互数目为 4. 利用 NetBeans6.9.1 编写的 JAVA 程序, 经 92 次迭代计算得到最少需要交互 17 次才能完成所有的消息传播. 最优的交互模式在表 1 中给出, 最优目标函数值随迭代次数的增加而减少的变化趋势如图 1(a) 所示.

如果将信息源增加到 16 个, 每次交互可传递的最大消息数为 5, 遗传的最大代次设定为 $T = 200$, 经 179 代遗传计算得到需要交互 82 次才能完成所有的消息传播. 最优个体的适应值随遗传代次的增加而减少的变化趋势如图 1(b) 所示.

从图 1 的最优遗传个体适应度变化趋势可以看出, 利用遗传算法求解单向最速网络消息传播问题是比较有效的. 但是, 遗传算法需要事先设定较多的

算法参数, 并且这些参数对随后的计算效率有较大影响, 因此, 只有对问题的特征有明确认识, 才能设计出高效的算法.

表 1 有 8 个信息源的最优交互模式
Tab.1 Best pattern of interchanging with 8 sources of messages

时间段	交互对与实际传递的消息
1	[(8,4),8],[(1,6),1],[(7,2),7],[(5,3),5]
2	[(1,8),1],[(6,2),6,1],[(4,3),8,4]
3	[(2,4),1,2,6,7],[(7,6),7],[(3,8),3,4,5]
4	[(8,5),1,3,4,8],[(3,6),3,4,5,8],[(2,7),1,2,6],[(4,1),2,6,7,8]
5	[(8,4),3,5],[(6,1),3,4,5],[(2,3),1,2,6,7]
6	[(2,8),2,6,7]
7	[(1,7),3,4,5,8],[(3,5),2,6,7],[(8,6),2],[(4,2),3,4,5,8]

注: 中括号中的小括号内是交互的两个信息源标号, 随后是该交互的实际传播消息标号.

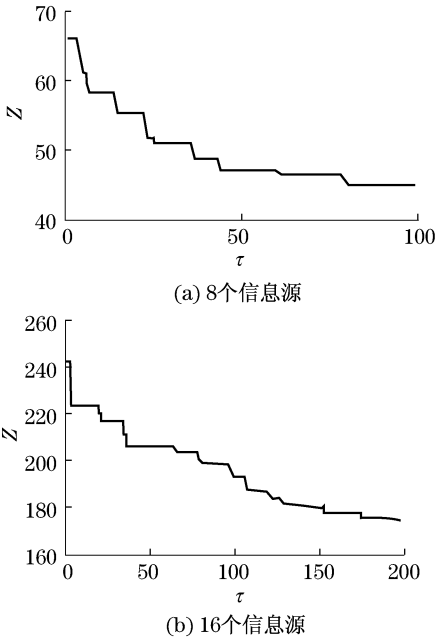


图 1 有 8 个和 16 个信息源的最优遗传个体的适应度变化趋势
Fig.1 Variation trend of the fitness of the best genetic individual with 8 and 16 sources of messages

4 结束语

通过对单向限量最速网络消息传播问题的建模和求解, 为相关的研究做一个铺垫. 网络消息传播问题的决策变量形式特殊、决策的时空影响模式特殊及问题的广泛适应性和强可塑性, 使得很难构造一

个该问题的普适模型. 本文给出的单向限量消息传播问题最优化模型提供了相关建模研究的基础. 利用问题特征设计的模拟进化算法能方便快捷地求解该问题, 同时具有较强的可塑性. 在此算法基础上可以开发更有效的问题族求解算法.

最速网络消息传播问题是一族问题的统称. 该问题族存在众多的扩展方向, 其中包括: a. 事先限定网络的信息交互通道的拓扑结构问题; b. 网络信息分布初态和终态的可变性问题; c. 信息传播过程中的消息遗忘问题; d. 多阶段的网络信息分布控制优化问题等. 对最速网络消息传播问题族的研究才刚刚开始, 存在许多基础理论问题亟待解决, 不仅要发掘问题的有效解决手段, 也要对问题的本质特征和规律进行深入的分析. 可以期待不久的将来最速网络消息传播问题族研究会有更多的成果出现, 也可以相信该问题族将会有广阔的实际应用前景.

参考文献:

- [1] 罗莹, 刘冰. 网络信息传播效果研究[J]. 情报科学, 2009, 27(9): 1487 - 1491.
- [2] 王慧军, 王有远, 胡振鹏. 网络信息传播管理研究[J]. 科技管理研究, 2010, 30(6): 140 - 143.
- [3] 孙丽县. 网络信息传播: 一种自组织复杂系统的分析研究[J]. 现代情报, 2007, 27(4): 81 - 84.
- [4] 金镇. 论网络信息传播的跨学科研究[J]. 现代情报, 2007, 27(4): 15 - 17.
- [5] 边霞, 米良. 遗传算法理论及其应用研究进展[J]. 计算机应用研究, 2010, 27(6): 2425 - 2434.
- [6] 沐士光. 遗传算法在网络优化问题中的研究与应用[J]. 计算机仿真, 2010, 27(4): 128 - 131.
- [7] 洪玮, 崔杜武. 函数优化的遗传算法策略优选[J]. 计算机工程与设计, 2010, 31(13): 3043 - 3046.
- [8] 贾东立, 郑国莘. 基于混沌和高斯局部优化的混合差分进化算法[J]. 控制与决策, 2010, 25(5): 899 - 902.
- [9] ALI M M, KAJEE-BAGDADI Z. A local exploration-based differential evolution algorithm for constrained global optimization[J]. Applied Mathematics and Computation, 2009, 208(1): 31 - 48.
- [10] 祝希路, 王柏. 一种基于社团划分的小生境遗传算法[J]. 控制与决策, 2010, 25(6): 1113 - 1116.
- [11] NEWMAN M E J, GIRVAN M. Finding and evaluating community structure in networks[J]. Physical Review E, 2004, 69(22): 1 - 15.
- [17] CRUZ J M, LIU Z. Modeling and analysis of the multi-period effects of social relationship on supply chain networks[J]. European Journal of Operational Research, 2011, 214(1): 39 - 52.
- [18] NAGURNEY A. Network Economics: A Variational Inequality Approach[M]. 2nd. Dordrecht: Kluwer Academic Publishers, 1999.
- [19] NAGURNEY A, DONG J. Supernetworks: Decision-Making for the Information Age[M]. Cheltenham: Edward Elgar Publishing Inc, 2002.
- [20] BAZARAA M S, SHERALI H D, SHETTY C M. Nonlinear Programming: Theory and Algorithms[M]. New York: John Wiley & Sons, 1993.
- [21] GABAY D, MOULIN H. On the uniqueness and stability of Nash equilibria in noncooperative games [C]// BEN-SOUSSAN A, KLEINDORFER P, TAPIERO C S. Applied Stochastic Control in Econometrics and Management Science. Amsterdam: 1980, 271 - 294.

(上接第 267 页)