文章编号: 1001-0920(2004)12-1332-05

多Agent协商行为的效用分析

叶 斌[1,2], 马忠贵[1], 曾广平[3], 尹怡欣[3], 涂序彦[3]
(1. 北京理工大学 自动控制系, 北京 100081; 2 太原理工大学 矿业工程学院, 山西 太原 030024; 3. 北京科技大学 信息工程学院, 北京 100083)

摘 要: 给出多Agent协商行为的一种统计模型以及效用函数的表达式, 从统计的角度分析了多Agent协商行为的行为效用, 并给出了相关参数的定性分析, 从而为更好地设计多Agent系统的协商组织规则和协商策略提供了效用依据

关键词: 分布式人工智能; 多Agent; 协商; 效用
中图分类号: TP18 **文献标识码**: A

Utility analysis of multi-agent negotiation behavior

YE Bin[1,2], *MA Zhong-gui*[1], *ZENG Guang-ping*[3], *YIN Yi-xin*[3], *TU Xu-yan*[3]
(1. Department of Automatic Control, Beijing Institute of Technology, Beijing 100081, China; 2. School of Mining Engineering, Taiyuan University of Technology, Taiyuan 030024, China; 3. School of Information Engineering, Beijing University of Science and Technology, Beijing 100083, China. Conrrespondent: YE Bin, E-mail: yebin@sina.com)

Abstract: A statistical model and a formula of utility function based on multi-agent negotiation behavior are presented. The utility of multi-agent negotiation behavior is analyzed from the statistical point of view. The parameters of multi-agent negotiation behavior are also analyzed. The results provid the foundation of utility analysis for designing the structure rules and negotiation strategy of multi-agent system.

Key words: distributed artificial intelligence; multi-agent; negotiation; utility

1 引 言

Agent技术在计算机领域的研究和应用源于20世纪70年代美国麻省理工学院研究人员开展的一系列关于分布式人工智能的研究 多Agent系统(MAS)[1]是指由多个Agent组成的系统, 用于解决单个Agent所不能解决的复杂问题, 由多个Agent协调合作形成问题求解网络 为使Agent之间合理高效地进行协作, Agent之间的协商[2]和协调机制成为多Agent系统研究的重点问题 通过协商, 多个Agent对其目标和资源等进行合理安排, 以调整各自的行为, 最大程度地实现各自目标

从效用理论[3]的角度看, 每个Agent都希望用最小的行为代价或花费来实现目标 协商过程主要依赖于Agent最大化其效用的内部目标和最小化其行为代价 那么, 从协商行为的角度看, 需要协商的MAS的组成规模是否有限制? 是否存在一个最优值? 其组织方式或组织规则[4,5]是否是随意的? 协商时间的长短对系统的稳定性有何影响? 为解决诸如此类的问题, 有必要从宏观角度对Agent的协商行为进行研究

2 问题描述

设有两组多Agent系统: $MAS_1 = A = \{A_i\}$,

收稿日期: 2004-03-22; **修回日期**: 2004-06-11.
基金项目: 国家自然科学基金资助项目(60375038); 国家863计划项目(863-511-944-019).
作者简介: 叶斌(1971—), 男, 山西平定人, 讲师, 博士生, 从事人工智能、智能信息处理等研究; 涂序彦(1935—), 男, 江西南昌人, 教授, 博士生导师, 从事人工智能及应用、人工生命等研究

$MAS_2 = B = \{B_c\}$, A 中的Agent随机到来与B 中的Agent进行协商完成任务, 然后离开系统. 如果A_i 到来时, 系统B 处于繁忙状态, 即没有空闲的B_c 与A_i 协商, 则A_i 将排队等待协商. 系统采用"先到先协商"的分配优先权规则执行协商活动. 协商行为过程如图 1 和图 2 所示.

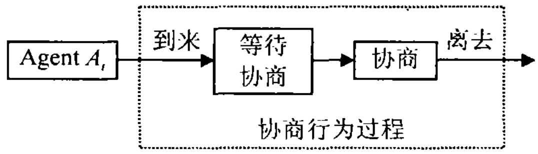

图 1　协商行为过程

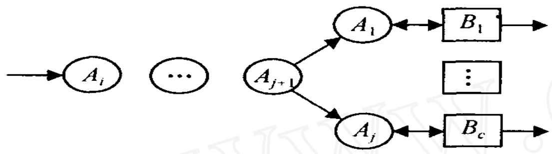

图 2　协商优先权分配规则

在Agent 应用领域中, 这类问题经常存在, 如: 电子商务系统的网上购物、网上付款等; 电子政务的网上办公; 网上合同的签定; Agent 对服务器资源的请求服务等. 诸如此类涉及到多Agent 之间交互、协商的问题均可抽象为上述问题.

3　问题分析

3.1　建立模型

为便于分析, 将上述问题的前提条件用严格的数学语言加以描述:

1) 在$[0, T]$ 时间区间, $[s, s+t)$ 内到来 k 个Agent $A_i(i = 1, \cdots, k)$ 与$[0, s)$ 内到来A_j 的情况无关, 即此二事件相互独立;

2) 在$[s, s+t)$ 内到来的A_i 数只与区间的长度 t 有关, 而与时间起点 s 无关, 即过程的统计规律不随时间的推移而改变;

3) 在同一瞬间到来两个或两个以上A_i 是不可能的, 即在充分小的时间间隔中, 最多只到来一个$A_i(i = 1)$.

下面考察上述条件下的A_i 和B_c 的协商行为过程. 因为A_i 是随机到来的, 若用$A_i(t)$ 表示在$[0, t)$ 内到达的A_i 总数, 则对于每个给定时刻 t, $A_i(t)$ 都是一个随机变量. 所以随机变量族$\{A_i(t) \mid t \in [0, T)\}$ 便构成一个随机过程.

于是, 在随机过程$\{A_i(t) \mid t \in [0, T)\}$ 中, 前述3 个条件可描述为:

① 对任一组 $t_1 < t_2 < \cdots < t_n (n \geq 3)$, 随机变量 $A_i(t_2) - A_i(t_1)$, $A_i(t_3) - A_i(t_2)$, $\cdots$, $A_i(t_n) - A_i(t_{n-1})$ 相互独立;

② 对于$[s, s+t) \subset [0, T)$, 总有

$$P(A_i(s+t) - A_i(s) = k) = P(A_i(t) - A_i(0) = k) = P(A_i(t) = k),$$

其中设

$$P(A_i(0) = 0) = 1, \quad \sum_{k=0}^{\infty} P(A_i(t) = k) = 1;$$

③ 令$\Psi(t) = \sum_{k=2}^{\infty} P(A_i(t) = k)$, 则$\lim_{t \to 0} \frac{\Psi(t)}{t} = 0$.

显然, 以上条件满足随机过程的独立增量性、平稳性和普通性, 是一个泊松过程.

3.2　问题求解

设$\{A_i(t)\}$ 为描述A_i 到达情况的随机过程, 以t_n 表示第 n 个A_i 到达的时刻. 则$T_n = t_n - t_{n-1}$ 为第 n 个Agent 与其前一个Agent 到达的时间间隔. 显然, $\{T_n\}$ 也是一族随机变量. 已经证明下述定理[6] 成立:

定理 1　Agent 到达过程$\{A_i(t)\}$ 为一个参数为 λ 的泊松过程, 其充要条件是: 相应的Agent 到达间隔$\{T_n\}$ 是一族相互独立的随机变量, 且每个随机变量都具有下列负指数分布:

$$P(T_n \leq t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0; \\ 0, & t < 0. \end{cases} \tag{1}$$

因为A_i 的到来是一个泊松过程(如前所述), 所以由定理 1 可知, $\{T_n\}$ 是负指数分布. 若A_i 到来时B_c 正处于繁忙状态, 则A_i 将等待协商. 设A_i 与B_c 的协商时间为$\{C_n\}$, 属负指数分布(由定理 1 可知), 则A_i 到达的时间间隔$\{T_n\}$ 的分布为式(1), A_i 与B_c 的协商时间$\{C_n\}$ 的分布为

$$P(C_n \leq t) = \begin{cases} 1 - e^{-\mu t}, & t \geq 0; \\ 0, & t < 0. \end{cases} \tag{2}$$

式中: λ 为Agent 的平均到达率, 即单位时间内平均到来的Agent 个数; μ 为平均协商率, 即单位时间内与B 协商的Agent(A_i) 的平均数目.

定义 1(生灭过程[7])　设在时刻 t 时 $\eta(t) = j$, 则在时刻 $t + \Delta t$ 时, $\eta(t + \Delta t) = j + 1$ 的概率为$\lambda_j \Delta t + 0(\Delta t)$, 其中$\lambda_j > 0$ 为与 t 无关的常数; 在时刻 $t + \Delta t$ 时, $\eta(t + \Delta t) = j - 1$ 的概率为$\mu_j \Delta_t + 0(\Delta t)$, 其中$\mu_j > 0$ 也是与 t 无关的常数; 在时刻 $t + \Delta t$ 时, $\eta(t + \Delta t)$ 为 I(随机变量 $\eta(t)$ 的取值集合为可列集 $I = \{0, 1, \cdots, m\}$) 中其他元素的概率均为$0(\Delta t)$. 满足上述条件的随机过程$\{\eta(t) \mid t \in [0, T)\}$ 称为生灭过

程

定义 2 (稳态概率) 令 $\pi_0 = 1$, $\pi_j = \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j}(j = 1, 2, \cdots)$, 并设生灭过程$\{\eta(t) \mid t \geq 0\}$的状态集合为 $I = \{0, 1, \cdots, m\}$ 或 $I = \{0, 1, \cdots\}$. 则当条件 $\sum_{j=0}^{\infty}\pi_j < +\infty$, $\sum_{j=0}^{\infty}\frac{1}{\lambda_j\pi_j} = +\infty$ 满足时, 对于 $\forall s > 0$ 和 $\forall j \in I, i \in I$, 都有下式成立:

$$\lim_{t\to\infty} P(\eta(t) = j) =$$

$$\lim_{t\to\infty} P(\eta(s + t) = j \mid \eta(s) = i) = p_i > 0.$$

当状态集 $I = \{0, 1, \cdots, m\}$ 时,

$$p_0 = \Big(\sum_{j=0}^{\infty}\pi_j\Big)^{-1}, p_j = \pi_j p_0 = \frac{\lambda_{j-1}}{\mu_j}p_{j-1},$$

$$j = 1, 2, \cdots, m. \tag{3}$$

当状态集 $I = \{0, 1, \cdots\}$ 时,

$$p_0 = \Big(\sum_{j=0}^{\infty}\pi_j\Big)^{-1}, p_j = \pi_j p_0 = \frac{\lambda_{j-1}}{\mu_j}p_{j-1},$$

$$j = 1, 2, \cdots. \tag{4}$$

其中 $p_j(j = 0, 1, \cdots)$ 为生灭过程的稳态概率, 即统计平衡时, 系统中具有 j 个 Agent 的概率.

由定义 1 可知, 若用 $\eta(t)$ 表示在时刻 t 系统内所拥有的 Agent 数(即 A_i), 则可将$\{\eta(t)\}$看作一个生灭过程, 其状态集为 $I = \{0, 1, \cdots\}$.

由于 Agent 的到来过程是具有独立增量性和平稳性的参数为 λ 的泊松流, 由式(1) 知, 在 Δt 时间范围内到达一个 Agent 的概率为

$$P(\eta(\Delta t) = 1) = \frac{\lambda\Delta t}{1!}e^{-\lambda\Delta t}. \tag{5}$$

将式(5) 展开为 Taylor 级数, 有

$$P(\eta(\Delta t) = 1) = \frac{\lambda\Delta t}{1!}e^{-\lambda\Delta t} =$$

$$\lambda\Delta t(1 - \lambda\Delta t + 0(\Delta t)) = \lambda\Delta t + 0(\Delta t). \tag{6}$$

同理, 在 Δt 内一个 A_i 与 B_c 协商结束的概率为 $\mu\Delta t + 0(\Delta t)$ (由式(2) 可得). 对整个系统而言, 在 Δt 时间内结束对一个 Agent 协商活动的概率为 $C\mu\Delta t + 0(\Delta t)$.

下面计算该生灭过程的稳态概率. 对于该生灭过程, 有

$$\lambda_j = \lambda, j = 0, 1, \cdots;$$

$$\mu_j = \begin{cases} j\mu, j = 1, 2, \cdots, C; \\ C\mu, j = C + 1, C + 2, \cdots. \end{cases} \tag{7}$$

式中: C 为 B 拥有的 Agent 个数; j 为在 t 时刻内系统到来的所有 Agent 的个数(包括等候协商和正在协商的). 由定义 2 可知

$$\pi_j = \begin{cases} \frac{C^j}{j!}\rho^j, 1 \leq j \leq C; \\ \frac{C^C}{C!}\rho^j, j > C; \end{cases} \quad \rho = \frac{\lambda}{C\mu}. \tag{8}$$

式中: ρ 为 B 中每个 Agent 的工作强度, 即单位时间内协商过程所占用的平均时间; $1/\mu$ 为每个 Agent 的平均协商时间.

将上述参数代入式(4), 有

$$\sum_{j=0}^{\infty}\pi_j = \sum_{j=0}^{C}\pi_j + \sum_{j=C+1}^{\infty}\frac{C^C}{C!}\rho^j, \tag{9}$$

$$\sum_{j=0}^{\infty}\frac{1}{\lambda_j\pi_j} = \sum_{j=0}^{C}\frac{1}{\lambda\pi_j} + \sum_{j=C+1}^{\infty}\frac{1}{\lambda\pi_j} =$$

$$\frac{1}{\lambda}\sum_{j=0}^{C}\frac{j!}{\rho^j C^j} + \frac{C!}{\lambda C^C}\sum_{j=C+1}^{\infty}\frac{1}{\rho^j}. \tag{10}$$

式中 $1/\lambda$ 为 Agent 到达协商系统的平均间隔. 可见, $\rho < 1$ 时, $\sum_{j=0}^{\infty}\pi_j < +\infty$, $\sum_{j=0}^{\infty}\frac{1}{\lambda_j\pi_j} = \infty$. 由定义 2 可知, 此时有稳态概率

$$p_0 = \Big(\sum_{j=0}^{\infty}\pi_j\Big)^{-1} =$$

$$\Big(\sum_{j=0}^{C-1}\frac{(C\rho)^j}{j!} + \frac{(C\rho)^C}{C!}\frac{1}{1-\rho}\Big)^{-1}, \tag{11}$$

$$p_j = \begin{cases} \frac{(C\rho)^j}{j!}p_0, 1 \leq j \leq C; \\ \frac{C^C\rho^j}{C!}p_0, j > C. \end{cases} \tag{12}$$

由此可得

$$A = \sum_{j=0}^{\infty}jp_j = C\rho + \frac{\rho}{(1-\rho)^2}p_C, \tag{13}$$

$$A_q = \sum_{j=0}^{\infty}jp_{C+j} = \frac{\rho}{(1-\rho)^2}p_C, \tag{14}$$

$$W = \frac{1}{\lambda}A = \frac{1}{\mu} + \frac{\rho}{(1-\rho)^2}p_C, \tag{15}$$

$$W_q = \frac{1}{\lambda}A_q = \frac{\rho}{\lambda(1-\rho)^2}p_C. \tag{16}$$

式中: A 为系统中所容纳 Agent(A_i) 的数量 A (包括等候协商和正在协商的 Agent 数量) 的期望值; A_q 为等候协商的 Agent 数量的期望值; W 为 Agent 从到达系统至协商完毕离去为止所花费时间的期望值; W_q 为 Agent 从到达系统至开始协商所花费时间的期望值.

4 效用分析

4.1 效用函数的构造

4.1.1 一般表述

在协商背景下, 效用函数表示行为的代价或内部目标实现的度量, “效用”在这里是对 Agent 的系

统开销和资源占用的度量 假设A gent 的效用函数为 $u(x)$, A gent 的基础价格(或静态价值) 为 G, A gent 的行为价格(或动态价值, 此处为协商行为价格) 为 $P(t)$, 是时间的函数 于是, A gent 协商行为的效用函数为

$$\mu(A_i) = G + \int^{w} P(t)\,\mathrm{d}t, \qquad (17)$$

式中 w 为每个A gent 从到达系统至协商完毕离去所花费的时间 将式(15) 和(16) 代入式(17), 可得平均效用函数为

$$\overline{\mu}(A_i) = G + \int^{W} P(t)\,\mathrm{d}t =$$
$$G + \int^{W_q} P(t)\,\mathrm{d}t + \int_{W_q}^{W} P(t)\,\mathrm{d}t =$$
$$G + \int^{\frac{\rho}{\lambda(1-\rho)^2}p_C} P(t)\,\mathrm{d}t +$$
$$\int_{\frac{\rho}{\lambda(1-\rho)^2}p_C}^{\frac{1}{\mu}+\frac{\rho}{\lambda(1-\rho)^2}p_C} P(t)\,\mathrm{d}t \qquad (18)$$

4 1 2　具体表述

下面以MAS 在电子商务网上购物的一种简单协商行为为背景, 说明式(17) 和(18) 中参数的实际含义

假设 $MAS_1 = A = \{A_i\}$ 和 $MAS_2 = B = \{B_c\}$ 分别表示购物客户和网站购物中心的多A gent 系统, 若以导引协调[8] 模式来设计协商系统(如图3 所示), 则 A_i 进入协商系统后的购物协商行为可简单描述为: A_i 向 B 广播购物请求, B 根据 A_i 的请求采取相应的组织规则和协商策略, 从而形成协商系统 $\{A_i, B_c\}$; A_i 被 B 中的安全A gent B_1 验证身份, B_1 与客户管理A gent B_2 协商给予 A_i 的购物优惠政策, 同时, A_i 就其对所购物品的喜好、质量等有关问题与导购A gent B_3 进行协商, B_3 再与采购A gent B_4 和收银A gent B_5 协商货源、运输方式及价格、付款方式等问题, 最后将协商结果返回 A_i, 购物结束

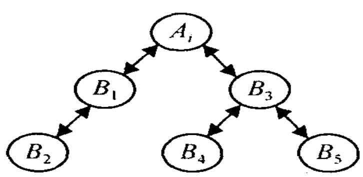

图 3　网上购物协商系统

在这里, 效用函数(17) 和(18) 的具体含义为:

$\mu(A_i)$ 为 A_i 从到达系统至购物结束所花费的系统开销和资源占用, 其值越小A gent 的效用越高;

G 为基础价格(或静态价值), 指 A_i 进入系统后即使不采取任何行为都会占用系统资源, 这里主要指所占的存储空间;

$P(t)$ 为行为价格(或动态价值), 此处为协商行为价格, 主要指 A_i 向 B 广播购物请求的代价与 B 中A gent 响应 A_i 的协商结果的代价之和 其表达式(推导过程略) 为

$$P(t) = p_1 \frac{k}{k-1}\left(k\frac{k^{h-1}-1}{k-1} - h + 1\right) +$$
$$p_2 \max\left\{\sum_{i=1}^{h-1}(i-1)w_i\right\}. \qquad (19)$$

式中: k 为图3 所示树的度, $k = 2$; h 为图3 所示树的深度, $h = 3$; W_i 为系统中A gent 的权值, 权值越大表示协商难度越大, 由A gent 收到请求信息后自我评估而确定; p_1 为单位时间内相邻A gent 之间传递请求信息所占用的系统资源量, 取统计平均值可简化计算; p_2 为单位时间内相邻A gent 之间进行协商并传递协商结果所占用系统资源量, 取统计平均值可简化计算

4 2　参数分析

由式(18) 可以看出, A gent 协商行为的效用与 μ, λ, p_C 有关, 进而与 A, A_q, W, W_q 有关 在设计协商系统的组织规则和协商策略时, 应充分考虑这些因素的影响, 使A gent 的协商行为的效用达到最优化

1) 如果已知 λ 和 μ, 便可计算出A gent 到达后等候协商的概率不需要等待即可协商的概率以及 B 中A gent 的工作强度, 从而合理利用系统资源:

① 由参数的含义可知, B 中A gent 的工作强度为 $\rho = \dfrac{\lambda}{C\mu}$;

② A gent 到达后, 若 $j \le C$, 即系统内到来的A gent 个数不超过 B 所拥有的A gent 个数, 则不需等待即可协商, 由式(11) 和(12) 可知, 其概率为 $p_j = \dfrac{(C\rho)^j}{j!}p_0$;

③ A gent 到达后, 若 $j > C$, 则必须等候协商, 由式(11) 和(12) 可知, 其概率为 $p_j = \dfrac{C^C\rho^j}{C!}p_0$

2) 如果已知 λ 和 μ, 便可求得 ρ, 进而求得A gent 从到达系统至协商完毕离去所花费时间的期望值, 即

$$W = \frac{1}{\mu} + \frac{C^{C-1}\rho^C}{\mu(1-\rho)^2 C!}p_0$$

3) 如果已知 λ 和 μ, 为使A gent 不至因等候时间过长而产生不耐烦情绪, 可以设定一个允许最长等候时间, 进而求得概率 p_{max}, 使A gent 到达后必须等候协商的概率不大于 p_{max}, 由此即可求得 B 中需

增加的Agent数量，从而合理设计系统容量.

4) 通过计算 p_C 可以分析系统的稳定性. 若 p_C 的值太大，甚至接近于1，则表明系统经常处于繁忙状态，随着时间的延续，系统会逐渐失稳.

5 实例分析

以一个电子商务网站的网上购物为例：Agent A_i 到一个电子商务网站，然后与网站的Agent B_c ($c = 2$) 进行协商购物. 下面对Agent B_c 的协商行为效用进行分析.

5.1 已知条件

1) 以2 min作为一个时段，依次记下各时段到来的Agent(A_i)数目，其平均值等于1.346次. 经分析知，这些数据的一个合适概型是：平均值为每时段(2 min)1.346的泊松分布.

2) 记下这些Agent A_i 与Agent B_c 协商所花费的时间，其平均值等于181.269 s. 经分析知，这些数据的一个合适概型是：平均值为3 min 12 s的负指数分布.

5.2 效用分析

由上述已知条件得：$c = 2, \lambda = 0.673, \mu = 0.331, \rho = 1.017$. 因为 $\rho > 1$，所以由式(9)和(10)可知，若将此协商系统维持很长一段时间，则系统中的Agent A_i 数目会趋于∞，从而使系统失去稳定，引起混乱.

为使系统达到稳定，可在系统中增加一个协商Agent B_3. 因此，$c = 3, \rho = 0.678$. 代入式(11)和(12)，可得

$$p_0 = 0.125, p_1 = 0.254,$$
$$p_2 = 0.259, p_3 = 0.362.$$

将上述参数代入式(15)和(16)，得

$$W = 4.72 \text{ min}, W_q = 1.70 \text{ min}.$$

计算结果表明：Agent B_c 的工作强度为0.678，协商系统的空闲率为0.125，每个Agent A_i 不必等候协商的概率为0.362. 将这些参数代入式(17)～(19)，分析Agent(A_i)的效用是有效的，而且Agent(A_i)可通过在不同的电子商务网站计算出不同的 $\mu(A_i)$，以优化决策自己的行为.

6 结 语

本文给出了多Agent协商行为的一种统计模型以及效用函数的表达式，从统计的角度分析了多Agent协商行为的行为效用，并给出了相关参数的定性分析，从而为更好地设计多Agent系统的协商组织规则和协商策略提供了效用依据.

在不同情况下，如何选择最优的协商模式，使Agent的协商行为的效用达到最优化？系统规模(即Agent的容量)是否存在最优值，如何确定？关于这些问题的深入探讨将另文给出.

参考文献(References)：

[1] 李海刚，吴启迪. 多Agent系统研究综述[J]. 同济大学学报，2003，31(6)：728-732.
(Li H G, Wu Q D. Summary on research of multi-agent system[J]. *J of Tongji University*, 2003, 31(6): 728-732.)

[2] 刘大有，杨鲲，陈建中. Agent研究现状与发展趋势[J]. 软件学报，2000，11(3)：315-321.
(Liu D Y, Yang K, Chen J Z. Agents: Present status and trends[J]. *J of Software*, 2000, 11(3): 315-321.)

[3] Stuart R, Peter N. *Artificial Intelligence: A Modern Approach*[M]. Beijing: Pearson Education North Asia Limited and People's Posts and Telecommunications Press, 2002: 473-484.

[4] Tu X Y, Tang T. Intelligent autonomous decentralized system[A]. *Proc of the 2nd Int Workshop of IEEE Computer Society on Autonomous Decentralized System*[C]. Beijing, 2002: 10-15.

[5] 张伟，石纯一. Agent组织的一种递归模型[J]. 软件学报，2002，13(11)：2149-2154.
(Zhang W, Shi C Y. A recursive model of agent organization[J]. *J of Software*, 2002, 13(11): 2149-2154.)

[6] 唐应辉，唐小我. 排队论——基础与应用[M]. 成都：电子科技大学出版社，2000：14-17.

[7] 魏国华，傅家良，周仲良. 实用运筹学[M]. 上海：复旦大学出版社，1987：389-390.

[8] 涂序彦. 大系统控制论[M]. 北京：国防工业出版社，1994：254-257.