

事件驱动的动态服务组合策略在线自适应优化

江 琦^{1,2}, 奚宏生², 殷保群²

(1. 合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009; 2. 中国科学技术大学 自动化系, 安徽 合肥 230027)

摘要: 针对复杂应用环境中网络新媒体服务系统的特点, 提出一种事件驱动的动态服务组合策略及其在线优化算法, 在保证各类业务服务质量(QoS)的同时, 提高系统资源的利用率. 通过定义不同类型的事件, 驱动服务组合的动态调整, 实现对各类业务QoS的保障和对业务需求变化的感知. 构建基于半Markov切换空间控制过程的系统分析模型, 利用模型的动态结构特点, 提出一种结合随机逼近和策略迭代的在线优化算法. 该算法不依赖系统参数信息, 对环境具有良好的自适应性. 仿真实验结果验证了算法的有效性.

关键词: 动态服务组合; 层次化Markov控制过程; 策略迭代; 服务覆盖网络; 优化算法

中图分类号: TP202 **文献标识码:** A

Online adaptive optimization for event-driven dynamic service composition

JIANG Qi^{1,2}, XI Hong-sheng², YIN Bao-qun²

(1. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei Anhui 230009, China;

2. Department of Automation, University of Science and Technology of China, Hefei Anhui 230027, China)

Abstract: An event-driven dynamic service composition strategy is presented for networked new media service systems to improve the resource utilization while simultaneously ensuring the quality of service(QoS) for multi-class services. An analytical model of semi-Markov switching state-space control processes is introduced to formulate the dynamic service composition problem. By utilizing the dynamic hierarchy of this model and the features of event-driven policy, an online optimization algorithm that combines stochastic approximation with the policy iteration is proposed. This algorithm is independent of any prior knowledge of system parameters and is with less computational cost. Simulation results demonstrate the effectiveness of the proposed algorithm.

Key words: dynamic service composition; hierarchical Markov control processes; policy iteration; service overlay networks; optimization algorithm

1 引言(Introduction)

随着信息技术的发展和应用需求的推动, 各类网络新媒体服务系统相继涌现. 所谓新媒体是指一切具有交互特征及数字化分布属性的数字媒体对象, 典型的新媒体业务如视频点播(VOD)、网络电视(IPTV)、远程教育、视频短信等. 新媒体业务通常具有数据处理量大、实时性要求高、服务请求数目多、连接时间长、多种服务质量(quality of service, QoS)指标要求等特点. 网络新媒体服务系统是面对日新月异的网络需求构建的多业务运行支撑平台, 它采用面向服务的设计, 具有功能组件和业务流程可复用重组、业务可快速生成发布、服务质量可监控等特点, 能够支持未来网络业务和技术迅速发展的需求.

网络新媒体服务系统处于复杂的网络应用环境, 各类业务需求具有高波动性, 用户的行为不可预见. 为有效提高系统对资源的整合能力、对业务需求变化的适应能力、对用户请求及其QoS的响应和保障能力, 通常需要在资源层、业务层和用户层对系统进行管理和控制. 通过资源的自适应部署、服务的动态组合、用户请求的调度和迁移等一系列管控策略的调度和执行, 网络新媒体服务系统构成一个支持多类型多尺度业务、业务流程灵活可重构、环境智能感知、可靠QoS保证、管控协同一体的智能化网络服务系统. 动态服务组合(dynamic service composition, DSC)对应于业务层面的管控, 在提高系统对业务需求变化的适应能力、保障各类业务的QoS以及提高系统资源的利用率上发挥着关键的作用.

收稿日期: 2010-01-28; 收修稿日期: 2010-11-29.

基金项目: 国家自然科学基金资助项目(60774038, 60904021, 61074033); 国家“863”计划资助项目(2008AA01A317); 中国博士后科学基金资助项目(20090450820); 博士学科点专项科研基金资助项目(20093402110019); 安徽省自然科学基金资助项目(11040606M142).

不同应用背景下的服务组合问题在多个层面得到了广泛的研究. 文献[1~3]研究基于语义的Web服务组合问题, 侧重服务发现、功能匹配及组合实现层面; 文献[4~6]采用AI-based Plan组合方法针对顺序执行的Web服务组合, 在提供业务QoS保证的同时, 降低组合成本; 文献[7~9]提出基于Markov决策过程的模型来处理Web服务组件的随机特性以提高组合的成功率和可靠性. 多媒体应用中的服务组合通常是多个逻辑上具有先后顺序的服务组件同时并行处理同一业务流, 由于服务的持续时间长, 在服务的持续过程中, 环境的动态变化对业务的QoS影响较大. 文献[10~12]将提供多媒体服务的服务覆盖网络(service overlay network, SON)转化为一个加权有向图, 进而应用Dijkstra最短路径算法^[10, 11], 或在分布式应用中通过有限步的探索^[12], 为到达的业务请求规划一条满足其QoS要求且能均衡负载的服务路径; 文献[13]考虑资源配置与业务需求匹配的DSC问题, 采用Markov决策过程来进行服务实例组合的优化, 但没有考虑服务持续过程中环境变化的影响.

网络新媒体服务系统中业务请求数量大、实时性要求高, 使得对各个具体的业务请求分别规划一条业务流程缺乏实际的可行性, 同时业务需求的高波动性又使得应用DSC来进行系统资源的适应性配置成为必要. 本文考虑复杂应用环境中网络新媒体系统在性能层面上的DSC问题, 提出一种事件驱动的DSC策略优化方法, 在多业务协同运行和业务需求波动的情况下, 提供各类业务QoS保证并有效提高系统资源的利用率. 通过定义不同类型的事件, 驱动服务组合的动态调整, 实现对各类业务QoS的保障和对业务需求变化的感知. 构建基于半Markov切换空间控制过程的系统分析模型, 精确刻画系统的层次化动态结构特征. 在此基础上提出一种结合随机逼近和策略迭代的在线DSC策略优化算法. 该算法充分利用事件驱动的策略特性^[14]和层次化的动态结构信息, 摆脱对系统和环境参数的依赖, 同时大幅降低计算的复杂度, 具有良好的自适应性和计算实时性. 通过仿真实验验证算法的有效性.

2 动态服务组合问题(Dynamic service composition problem)

网络新媒体服务系统是一个支持多类型业务服务的管控协同运行支撑平台. 系统将支持各类业务所需的功能分解为若干类具有适当颗粒度的、定义良好的服务功能, 并将分布于物理承载网络中的软硬件资源和内容通过资源配置构成相应类型的功能节点. 系统依托这些公共可复用的功能节点, 通过服务的组合灵活组建具有不同服务容量和QoS保证的多种类型的业务流程来支持相应类型的业务服务.

DSC即随着业务需求的变化(用户对不同种类业务的波动性需求、新业务需求的出现)和系统自身状态的改变(如网络部分链路的拥塞、某个节点的失效或部分失效)进行业务流程的重组. 其实质是根据业务需求的变化情况对系统可用资源在各类业务之间进行动态配置, 以适应业务需求的高波动, 为各类业务提供可靠的QoS保证. DSC通过与环境的频繁交互, 使得服务的组合对系统自身和环境的变化具有感知能力和自适应性.

网络新媒体服务系统通过资源部署和服务组合构成一个SON. 资源部署决定了SON节点(node, 即功能节点)的位置分布, 服务组合决定了节点间的连接(link). 在资源部署给定情况下, SON中各类功能节点的数量、服务容量以及位置分布已知, 节点间连接的网络链路带宽、时延等指标可测. 根据系统所需支持业务类型的功能性要求和QoS指标, 通过业务模版将业务流程映射到SON, 构成相应的业务链路. 一条业务链路对应一条服务路径(service path), 由若干功能节点通过网络链接按一定的逻辑顺序组合而成, 各节点和链接配置在此服务路径上的服务容量和链接带宽也预先确定. 同一种类型业务可以有多个服务路径, 不同类型的业务链路可以复用相同的功能节点. 同时满足系统资源约束的所有类型的业务链路构成一种服务组合方式, 所有可行的组合方式构成服务组合的可选集合. DSC根据业务需求和可用资源的变化情况, 选取其中的一种方式进行服务的组合, 从而确定了系统对各类业务所能提供的业务链路的数量和服务容量. 在此基础上, 各种类型业务的用户请求通过Portal调度至相应的业务链路上以提供连续的服务.

业务需求的动态变化以及网络链接和节点的状况变化驱动着服务的动态组合. 驱动服务动态组合的事件有: 当某种业务的流量增大, 使得当前的服务组合不能满足其容量需求, 或使得处于同一业务流程中其他尺度业务的容量需求不能满足; 某个节点的失效或部分失效; 组成业务链路的网络链接时延增大, 导致业务的QoS要求不能满足. 性能层次上的DSC问题即在给定的资源部署下, 从可选的服务组合方式中, 根据动态的业务需求和系统的运行情况适应性地选取组合方式进行服务的组合, 在保证各类业务QoS的同时, 使得系统的资源利用率(或等价地, 系统的收益)最大.

服务组合的一种方式, 决定了系统的一种动态运行规律. DSC驱动系统动态在多种运行规律之间进行切换, 使得系统的动态演化呈现出典型的层次化特征. 系统中用户请求的到达和服务的持续时间具有随机性, 网络链路的时延变化和节点失效情况的发生同样具有随机性. 在多媒体网络应用环境中,

这些随机变量通常满足或近似满足Poisson分布^[15]. 在DSC的驱动下, 系统的动态演化可以用一个两层级的半Markov切换空间控制过程来描述, 从而将DSC问题转化为一个随机优化问题.

3 系统分析模型(System analytical model)

考虑一个网络新媒体服务系统, 由 n_N 个功能节点构成, 为 n_I 类业务提供服务. 业务类型以 i 标记, 其集合为 I ; 节点类型用 j 表示, 其集合为 J ; 服务组合方式标记为 k , 所有可选组合方式构成集合 K ; 服务路径用 l 表示, 在组合方式 k 下的集合为 L^k , 其中提供 i 类业务服务的路径集合记为 L_i^k ; 服务节点用 n 表示, 所有服务节点构成集合 N , j 类节点构成的集合记为 N_j , 构成服务路径 l 的节点集合表示为 N_l ; 节点 n 的服务容量记为 C_n , 服务1个 i 类业务需占用 j 类节点的服务容量以 c_{ij} 表示; 服务路径 l 的服务容量 C_l . 对于 $\forall k \in K$, 有

$$\sum_{i \in I, l \in L_i^k} C_l \cdot c_{ij} \cdot I(n \in N_l \cap N_j) \leq C_n, n \in N,$$

其中 $I(\cdot)$ 为示性函数.

假设 i 类业务请求的到达服从到达率为 λ_i 的Poisson分布, 服务持续时间满足均值为 $\frac{1}{\mu_i}$ 的指数分布; 节点 n 失效和业务链路 l 的QoS降低的事件发生分别服从参数为 λ_n, λ_l 的Poisson分布, 恢复时间满足均值为 $\frac{1}{\mu_n}, \frac{1}{\mu_l}$ 的指数分布.

以 m_l 表示服务路径 l 中正在服务的业务请求的个数, 则在组合方式 k 下系统的状态可以表示为 $s = (m_l, l \in L^k)$, 相应地系统的状态空间为

$$S_k = \{(m_l, l \in L^k) \mid m_l \leq C_l, l \in L_i^k, i \in I\},$$

为简化表述, 这里没有考虑网络链接带宽的限制.

采用基于负载感知的用户请求调度策略, 即将新到达的用户请求分配至同种业务类型的服务路径中可用容量最大的那条. 对应于服务组合方式 k , 系统的动态演化规律可以用一个半Markov过程(连续时间Markov过程) M_k 来描述:

$$\{X_t, t \geq 0; S_k, A_k\},$$

其中转移速率矩阵 $A_k = [a_{ss'}^k]$, $s, s' \in S_k$, 其元素

$$a_{ss'}^k = \begin{cases} \lambda_i, & \dot{s} = s^{m_{l'}+1}, \\ m_{l'} \mu_i, & \dot{s} = s^{m_{l'}-1}, \\ -\sum_{\dot{s} \neq s} a_{s\dot{s}}, & \dot{s} = s, \\ 0, & \text{其他,} \end{cases}$$

式中:

$$s^{m_{l'}+1} = \{m_{l'}+1, l' = \arg \min_{l \in L_i^k} c_{al}; m_l, l \neq l', l \in L^k\} \in S^k,$$

$$s^{m_{l'}-1} = \{m_{l'}-1, l' \in L_i^k; m_l, l \neq l', l \in L^k\} \in S^k, \\ c_{al} = C_l - m_l.$$

定义当一个 i 类业务请求到达时, 当前服务组合情况下, 没有相应的服务容量来接纳的状况发生为事件 $e_i (i \in I)$, 业务链路 l 的QoS降低(链路时延超过业务要求、节点 $n \in N_l$ 失效或部分失效)或恢复分别为事件 $e_l, e'_l (l \in L)$, 以 $e_r (r = 1, 2, \dots, n_E)$ 来标记这些具体的事件, 事件集合 $E = \{e_r, r = 1, 2, \dots, n_E\}$.

控制行动为进行服务组合的重配置, 即改变服务的组合方式. 将服务组合方式从 k 切换至 k' 的控制行动记为 $d_{kk'}, k, k' \in K$, 可选行动集合 $D = \{d_{kk'}, k, k' \in K\}$.

控制行动由事件触发. 随机型DSC策略是事件集至行动集上概率分布的映射 $\mathcal{L}: E \rightarrow D^{[0,1]}$. 以 $\mathcal{L}_{d_{kk'}}(e_r)$ 表示事件 $e_r \in E$ 发生时选取行动 $d_{kk'} \in D$ 的概率. 随机型策略空间

$$\Pi = \{\mathcal{L} : 0 \leq \mathcal{L}_{d_{kk'}}(e_r) \leq 1,$$

$$\sum_{d_{kk'} \in D} \mathcal{L}_{d_{kk'}}(e_r) = 1, e_r \in E, d_{kk'} \in D\}.$$

当系统处于状态 s^k 时, 事件 e_r 发生, 采取行动 $d_{kk'}$, 系统下一个转移到的状态记为 $\vec{e}_{rs}^{kk'}$,

$$\vec{S}_{s^k} = \{\vec{e}_{rs}^{kk'}, e_r \in E, d_{kk'} \in D\}.$$

在策略 $\mathcal{L}$ 驱动下, 系统状态的演化规律可以用一个在多个半Markov过程之间切换构成的半Markov切换空间控制过程来描述:

$$\{X_t, D_t, t \geq 0; (M_k, k \in K), E, D\}.$$

定义性能函数 $f_r: \bigcup_{k \in K} S_k \rightarrow \mathbb{R}$, 表示系统处于状态 $s \in \bigcup_{k \in K} S_k$ 时服务各类业务所获得的收益率; 代价函数 $f_o: D \rightarrow \mathbb{R}$ 为采取切换行动 $d_{kk'} \in D$ 所对应的系统资源消耗.

考虑平均准则下的性能测度

$$\eta^{\mathcal{L}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (f_r(X_t) + f_o(D_t)) dt.$$

在此模型下, DSC问题为寻找一个最优策略 $\mathcal{L}^*$, 根据各类事件的发生情况, 对服务组合方式进行动态的调整, 在保证各类业务QoS的同时, 使得系统运行的收益最大 $\mathcal{L}^* \in \arg \max_{\mathcal{L} \in \Pi} \eta^{\mathcal{L}}$.

4 在线自适应策略优化(Optimization of the online adaptive strategy)

通过构建半Markov切换空间控制过程模型, DSC问题转换为一个无约束的随机优化问题. 当系统参数已知或能够通过仿真的方法预先获取时, 可应用线性规划等数值计算方法进行DSC策略的离线优化. 由于网络新媒体服务系统DSC的实际应用

环境复杂,系统参数通常难以预先获取,而系统巨大的状态空间也使得数值计算方法面临着“维数灾难”问题.针对这样的应用情况,提出一种结合性能势在线估计与策略迭代的在线自适应优化算法.

上述半Markov切换空间控制过程是一个具有特殊动态结构的半Markov控制过程^[16]

$$\{X_t, D_t, t \geq 0; S, A^{\mathcal{L}}, E, D, (f_r, f_o), \eta^{\mathcal{L}}\},$$

其状态空间 $S = \bigcup_{k \in K} \{k\} \times S_k$, 以 $s^k, s \in S_k, k \in K$

表示其中的状态;转移速率矩阵 $A^{\mathcal{L}} = [a_{s^k \dot{s}^k}^{\mathcal{L}}]$, $s^k, \dot{s}^k \in S$, 其中元素

$$a_{s^k \dot{s}^k}^{\mathcal{L}} = \begin{cases} \lambda_{s^k + e_r} \mathcal{L}_{d_{kk'}}(e_r), & \dot{s}^k = \vec{e}_{rs}^{kk'}, \in \vec{S}_{s^k}, \\ a_{s^k \dot{s}^k}^k, & k = k', \dot{s}^k \notin \vec{S}_{s^k}, \\ 0, & \text{其他,} \end{cases} \quad (1)$$

式中 $\lambda_{s^k + e_r}$ 为系统处于状态 s^k 时,事件 e_r 发生的速率.

转移速率矩阵 $A^{\mathcal{L}}$ 包含了过程的层次化动态结构信息并体现了事件驱动的策略特性,其主对角线为对应于 $S_k (k \in K)$ 的方块阵,其他非零元素皆与策略相关,在包含策略的元素中策略参数以乘积的形式出现.这些特性在优化算法中充分加以利用,可以有效摆脱算法对系统参数的依赖,并大幅减小计算量.

对于 $\forall \mathcal{L} \in \Pi$, 上述半Markov过程不可约,系统的稳态分布 $\mathbf{p}^{\mathcal{L}} = (p_{s^k}^{\mathcal{L}}, s^k \in S)$ 存在且唯一,满足平衡方程

$$\mathbf{p}^{\mathcal{L}} A^{\mathcal{L}} = 0, \mathbf{p}^{\mathcal{L}} \mathbf{e} = 1, A^{\mathcal{L}} \mathbf{e} = 0,$$

其中 $\mathbf{e} = (1, 1, \dots, 1)^T$.

系统的性能测度

$$\begin{aligned} \eta^{\mathcal{L}} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (f_r(X_t) + f_o(D_t)) dt = \\ &= \sum_{s^k \in S} p_{s^k}^{\mathcal{L}} f_r(s^k) + \sum_{s^k \in S} p_{s^k}^{\mathcal{L}} \sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}} f_o(d_{kk'}) = \\ &= \sum_{s^k \in S} p_{s^k}^{\mathcal{L}} f_{s^k}^{\mathcal{L}}, \end{aligned}$$

其中 $f_{s^k}^{\mathcal{L}} = f_r(s^k) + \sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}} f_o(d_{kk'})$.

设 $\{X_t, D_t, t \geq 0\}$ 为策略 $\mathcal{L}$ 驱动下上述半Markov控制过程的一条样本轨道,定义状态 $s^k \in S$ 的性能势^[17]

$$g_{s^k}^{\mathcal{L}} = E\left[\int_0^\infty (f_r(X_t) + f_o(D_t) - \eta^{\mathcal{L}}) dt \mid X_0 = s^k\right].$$

性能势向量 $\mathbf{g}^{\mathcal{L}} = (g_{s^k}^{\mathcal{L}}, s^k \in S)^T$ 满足如下Poisson方程:

$$-\mathbf{f}^{\mathcal{L}} + \eta^{\mathcal{L}} \mathbf{e} = A^{\mathcal{L}} \mathbf{g}^{\mathcal{L}},$$

其中 $\mathbf{f}^{\mathcal{L}} = (f_{s^k}^{\mathcal{L}}, s^k \in S)^T$.

设 $A^{\mathcal{L}'}$ 为策略 $\mathcal{L}' \in \Pi$ 诱导的另一个转移速率矩阵, $\mathbf{p}^{\mathcal{L}'}, \eta^{\mathcal{L}'}$ 是对应的稳态概率和平均性能测度,将

Poisson方程的两边左乘 $\mathbf{p}^{\mathcal{L}'}$ 并运用 $\mathbf{p}^{\mathcal{L}'} A^{\mathcal{L}'} = 0$, 得到性能差公式

$$\eta^{\mathcal{L}'} - \eta^{\mathcal{L}} = \mathbf{p}^{\mathcal{L}'} ((A^{\mathcal{L}'} - A^{\mathcal{L}}) \mathbf{g}^{\mathcal{L}} + (\mathbf{f}^{\mathcal{L}'} - \mathbf{f}^{\mathcal{L}})). \quad (2)$$

根据性能差公式可以基于一个策略下的性能势比较两个策略的优劣,有以下引理.

引理 1 策略 $\mathcal{L}$ 是最优的,当且仅当对于所有的 $\mathcal{L}' \in \Pi$, 有

$$A^{\mathcal{L}} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}} \geq A^{\mathcal{L}'} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}'}. \quad (3)$$

证 首先证明“当且”.若式(3)成立,因 $\mathbf{p}^{\mathcal{L}} > 0$, 由式(2),有 $\eta^{\mathcal{L}} \geq \eta^{\mathcal{L}'}$, 即 $\mathcal{L}$ 是最优策略.

其次证明“仅当”.假设存在一个策略 $\mathcal{L}'$, 使得式(3)不成立,则至少存在一个分量,满足

$$\sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}} g_{\dot{s}^k}^{\mathcal{L}} + f_{s^k}^{\mathcal{L}} < \sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}'} g_{\dot{s}^k}^{\mathcal{L}'} + f_{s^k}^{\mathcal{L}'},$$

即有

$$\begin{aligned} \sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}} (g_{\dot{s}^k}^{\mathcal{L}'} + f_o(d_{kk'})) &< \\ \sum_{\dot{s}^k \in S} a_{s^k \dot{s}^k}^{\mathcal{L}'} (g_{\dot{s}^k}^{\mathcal{L}'} + f_o(d_{kk'})). \end{aligned}$$

结合式(1),则存在 $e_r \in E$, 使得

$$\begin{aligned} \sum_{\dot{s}^k \in \vec{S}_{s^k}} \lambda_{s^k + e_r} \mathcal{L}_{d_{kk'}}(e_r) (g_{\dot{s}^k}^{\mathcal{L}'} + f_o(d_{kk'})) &< \\ \sum_{\dot{s}^k \in \vec{S}_{s^k}} \lambda_{s^k + e_r} \mathcal{L}'_{d_{kk'}}(e_r) (g_{\dot{s}^k}^{\mathcal{L}'} + f_o(d_{kk'})), \end{aligned}$$

构造策略 $\mathcal{L}'' : \mathcal{L}''(e_r) = \mathcal{L}'(e_r), \mathcal{L}''(e_n) = \mathcal{L}(e_n), e_n \in E, e_n \neq e_r$. 则

$$A^{\mathcal{L}''} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}''} \geq A^{\mathcal{L}} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}}.$$

由式(2),有 $\eta^{\mathcal{L}''} > \eta^{\mathcal{L}}$, 这与 $\mathcal{L}$ 是最优策略的假设相矛盾. 证毕.

定义 $\omega_{e_{rs}^{kk'}}^{\mathcal{L}} = g_{\dot{s}^k}^{\mathcal{L}'} + f_o(d_{kk'})$, 其中 $\dot{s}^k = \vec{e}_{rs}^{kk'}$.

定理 1 对所有的 $e_r \in E$, 若

$$\omega_{e_{rs}^{kk'}}^{\mathcal{L}} = \max_{d_{kk''} \in D} \omega_{e_{rs}^{kk''}}^{\mathcal{L}},$$

置 $\mathcal{L}'_{d_{kk'}}(e_r) = 1, \mathcal{L}'_{d_{kk''}}(e_r) = 0, k'' \neq k'$, 则 $\eta^{\mathcal{L}'} \geq \eta^{\mathcal{L}}$. 并且,若 $\mathcal{L}' = \mathcal{L}$, 则 $\mathcal{L}$ 是最优策略.

证 由式(1), $A^{\mathcal{L}'} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}'} - (A^{\mathcal{L}} \mathbf{g}^{\mathcal{L}} + \mathbf{f}^{\mathcal{L}})$ 的第 $s^k \in S$ 分量

$$\begin{aligned} &\sum_{\dot{s}^k \in S} (a_{s^k \dot{s}^k}^{\mathcal{L}'} - a_{s^k \dot{s}^k}^{\mathcal{L}}) g_{\dot{s}^k}^{\mathcal{L}} + (f_{s^k}^{\mathcal{L}'} - f_{s^k}^{\mathcal{L}}) = \\ &\sum_{\dot{s}^k \in S} (a_{s^k \dot{s}^k}^{\mathcal{L}'} - a_{s^k \dot{s}^k}^{\mathcal{L}}) g_{\dot{s}^k}^{\mathcal{L}} + \\ &\sum_{\dot{s}^k \in S} (a_{s^k \dot{s}^k}^{\mathcal{L}'} - a_{s^k \dot{s}^k}^{\mathcal{L}}) f_o(d_{kk'}) = \\ &\sum_{\dot{s}^k \in S} (a_{s^k \dot{s}^k}^{\mathcal{L}'} - a_{s^k \dot{s}^k}^{\mathcal{L}}) (g_{\dot{s}^k}^{\mathcal{L}} + f_o(d_{kk'})) = \\ &\sum_{e_r \in E} \lambda_{s^k + e_r} (\omega_{e_{rs}^{kk'}}^{\mathcal{L}} - \sum_{d_{kk''} \in D} \mathcal{L}_{d_{kk''}}(e_r) \omega_{e_{rs}^{kk''}}^{\mathcal{L}}), \end{aligned}$$

由于 $\lambda_{s^k+e_r} \geq 0$, $\sum_{d_{kk'} \in D} \mathcal{L}_{d_{kk'}}(e_r) = 1$, $e_r \in E$, 故上式大于等于 0, 根据引理 1 可得期望结论。证毕。

定理 1 表明, 最优 DSC 策略可在确定型策略空间

$$\Pi_D = \{\mathcal{L} : \mathcal{L}_{d_{kk'}}(e_r) = 1 \text{ 或 } 0, d_{kk'} \in D, \sum_{d_{kk'} \in D} \mathcal{L}_{d_{kk'}}(e_r) = 1, e_r \in E\}$$

取得, 因而可限定于在 Π_D 中进行 DSC 策略寻优, 适用基于策略迭代的优化算法。

根据定理 1, 基于 $\omega_{e_{rs}}^{kk'}$ 值可以进行策略的改进及最优策略的判定, 结合性能势的在线估计, 有如下策略迭代优化算法。

算法 1 在线自适应 DSC 策略优化算法。

Step 1 置 $m = 1$, $n = 1$, $s_0 = s^*$, 并对于所有的 $e_r \in E$, 置 $\omega_{e_{rs}}^{kk'} = 0$; 选定一个初始策略 $\mathcal{L}^1 \in \Pi$ 。

Step 2 在 $\mathcal{L}^m$ 下运行系统, 观测状态的演化; 记录第 m 周期中的第 $(n+1)$ 次状态转移时刻 t_m^{n+1} 和转移到的状态 $s_{t_m^{n+1}}$, 计算 $s_{t_m^n}$ 的逗留时间 $T_m^n = t_m^{n+1} - t_m^n$ 。

Step 3 若 $s_{t_m^{n+1}} = s^*$, 令 $T_m = t_m^{n+1} - t_m^1$, 对于所有的 $s_{t_m^n} \in \tilde{S}_{s^k}$, $s^k \in S$, 计算性能势估计值 $\hat{g}_{s_{t_m^n}}^{\mathcal{L}^m}$:

$$\begin{aligned} \hat{g}_{s_{t_m^n}}^{\mathcal{L}^m} &= \sum_{k=n}^{n_m} [(f_r(s_{t_m^k}) - \hat{\eta}^{\mathcal{L}^m})T_m^k + f_o(d_{t_m^{k+1}})], \\ \hat{\eta}^{\mathcal{L}^m} &= \frac{1}{T_m} \sum_{n=1}^{n_m} (f_r(s_{t_m^n})T_m^n + f_o(d_{t_m^{n+1}})). \end{aligned}$$

否则, $n = n + 1$, 转 Step 2。

Step 4 对每一 $e_r \in E$ 和 $d_{kk'} \in D$, 若 $\dot{s}^{kk'} = \tilde{e}_{rs}^{kk'} \in \tilde{S}_{s^k}$ 在第 m 再生周期出现, 计算 $\hat{\omega}_{e_{rs}}^{kk'} = \hat{g}_{\dot{s}^{kk'}}^{\mathcal{L}^m} + f_o(d_{kk'})$; 否则 $\hat{\omega}_{e_{rs}}^{kk'} = \omega_{e_{rs}}^{kk'-1}$. 更新 $\omega_{e_{rs}}^{kk'}$, $e_r \in E$, $d_{kk'} \in D$,

$$\omega_{e_{rs}}^{kk'} = \omega_{e_{rs}}^{kk'-1} + \frac{1}{m} (\hat{\omega}_{e_{rs}}^{kk'} - \omega_{e_{rs}}^{kk'-1}). \quad (4)$$

Step 5 对每一 $e_r \in E$, 若 $\omega_{e_{rs}}^{kk'} = \max_{d_{kk'} \in D} \omega_{e_{rs}}^{kk''}$, 置 $\mathcal{L}_{d_{kk'}}(e_r) = 1$ 及 $\mathcal{L}_{d_{kk''}}(e_r) = 0$, $k'' \neq k'$. 若

$$\begin{aligned} \sum_{e_r \in E} \sum_{d_{kk'} \in D} \mathcal{L}_{d_{kk'}}(e_r) \omega_{e_{rs}}^{kk'} &> \\ \sum_{e_r \in E} \sum_{d_{kk'} \in D} \mathcal{L}_{d_{kk'}}^m(e_r) \omega_{e_{rs}}^{kk'}, \end{aligned} \quad (5)$$

则置 $\mathcal{L}^{m+1} = \mathcal{L}$; 否则, 置 $\mathcal{L}^{m+1} = \mathcal{L}^m$ 。

Step 6 置 $t_{m+1}^1 = t_m^{n+1}$, $s_{t_{m+1}^1} = s^*$, 以及 $m = m + 1$, $n = 1$, 转 Step 2。

算法 1 基于系统运行这一样本轨道在线进行策略优化, 其输入为: 当前策略 $\mathcal{L}^m$ 、 $|S|$ 个状态性能函数值、 $|D|$ 个行动代价函数值以及 n_m 个状态转移时刻 t_m^n 和转移到的状态 $s_{t_m^n}$ (n_m 为第 m 个再生周期中状态转移的次数, 符号 $|\cdot|$ 表示集合中元素的个数), 输出为改进的策略 $\mathcal{L}^{m+1}$. 由计算步骤可知, 需要进行 $9n_m + |E||D|$ 次算术运算, 所需的存储空间

为 $6n_m + |S| + |D| + |E||D|$, 故算法 1 的计算复杂度为 $O(n_m)$ 。

算法 1 进行策略评估 (Step 4) 和策略改进 (Step 5) 时都不依赖于系统参数的信息 (如转移速率矩阵), 具有良好的环境自适应性; 同时仅需计算与事件相关的性能势估计值 (Step 3), 相对于基于状态的策略迭代算法^[18,19], 其计算量大大减小, 提高了计算的实时性。

定理 2 存在一个 M (依赖于样本轨道), 使得所有由算法 1 产生的 $\mathcal{L}^m$ ($m \geq M$) 是最优策略, 即算法 1 以概率 1 收敛:

$$\lim_{m \rightarrow \infty} \mathcal{L}^m \in \arg \max_{\mathcal{L} \in \Pi} \eta^{\mathcal{L}}, \text{ w.p.1.}$$

证 记 $\mathcal{B}$ 为算法 1 无限次访问的策略集合, 设 $\mathcal{L} \in \mathcal{B}$ 使得 $\eta^{\mathcal{L}} = \max_{\mathcal{L}' \in \mathcal{B}} \eta^{\mathcal{L}'}$ 。

首先证明 $\mathcal{B} = \{\mathcal{L}\}$. 用反证法, 假设 $\mathcal{B} \neq \{\mathcal{L}\}$, 则存在一个 $\mathcal{L}' \in \mathcal{B}$, 从而算法从 $\mathcal{L}$ 无穷多次切换到 $\mathcal{L}'$. 由 $\eta^{\mathcal{L}'} \leq \eta^{\mathcal{L}}$ 以及式 (2), 有

$$\sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \left(\sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) g_{\dot{s}^{kk'}}^{\mathcal{L}} + (f_{s^k}^{\mathcal{L}'} - f_{s^k}^{\mathcal{L}}) \right) \leq 0,$$

或等价地,

$$\sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \omega_{e_{rs}}^{\mathcal{L}} \leq 0.$$

记

$$\begin{aligned} \sigma_0 &= \inf\{m \geq 0, \mathcal{L}^m = \mathcal{L}\}, \\ \tau_u &= \inf\{m > \sigma_u, \mathcal{L}^m \neq \mathcal{L}\}, \\ \sigma_{u+1} &= \inf\{m > \tau_u, \mathcal{L}^m = \mathcal{L}\}. \end{aligned}$$

在 τ_u 时刻, 策略从 $\mathcal{L}$ 切换至其他策略, 在此期间切换至 $\mathcal{L}'$ 无穷多次. 在第 σ_u 时刻, 策略返回至 $\mathcal{L}$. 根据算法 1 的 Step 5, 并且注意到对于所有的 s^k 和 $\mathcal{L}' \in \Pi$, $p_{s^k}^{\mathcal{L}'} > 0$, 存在 $C > 0$, 不等式

$$\sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \omega_{e_{rs}}^{\tau_u} \geq C, \quad (6)$$

$$\sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \omega_{e_{rs}}^{\sigma_u} \leq -C \quad (7)$$

成立无穷多次. 由式 (4), 有

$$\omega_{e_{rs}}^{\tau_u} = \frac{\sigma_u}{\tau_u} \omega_{e_{rs}}^{\sigma_u} + \frac{1}{\tau_u} \sum_{v=\sigma_u}^{\tau_u-1} \hat{\omega}_{e_{rs}}^{v+1}, \quad (8)$$

将式 (8) 代入式 (6), 并运用式 (7), 得到

$$\begin{aligned} \sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \frac{1}{\tau_u} \sum_{v=\sigma_u}^{\tau_u-1} \hat{\omega}_{e_{rs}}^{v+1} &\geq \\ - \sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \frac{\sigma_u}{\tau_u} \omega_{e_{rs}}^{\sigma_u} + C &\geq C. \end{aligned}$$

这表明下面的不等式成立无穷多次:

$$\sum_{s^k \in S} p_{s^k}^{\mathcal{L}'} \sum_{\dot{s}^{kk'} \in \tilde{S}_{s^k}} (a_{s^k \dot{s}^{kk'}}^{\mathcal{L}'} - a_{s^k \dot{s}^{kk'}}^{\mathcal{L}}) \sum_{v=\sigma_u}^{\tau_u-1} \hat{\omega}_{e_{rs}}^{v+1} \geq C \tau_u. \quad (9)$$

由于 $\tau_u - \sigma_u < \infty$, $\omega_{e_{rs}^{kk'}}^{\mathcal{L}^{v+1}} < \infty$ 及 $E[\hat{\omega}_{e_{rs}^{kk'}}^{v+1}] = \omega_{e_{rs}^{kk'}}^{\mathcal{L}^{v+1}}$, 有

$$P\left\{\sum_{v=\sigma_u}^{\tau_u-1} \hat{\omega}_{e_{rs}^{kk'}}^{v+1} \geq C'\tau_u, \text{i.o.}\right\} = 0.$$

结合式(1), 表明不等式(9)不可能成立无穷多次. 这同假设矛盾, 故 $\mathcal{B} = \{\mathcal{L}\}$, 且存在一个 M 使得对于 $m \geq M$, $\mathcal{L}^m = \mathcal{L}$.

其次证明 $\mathcal{L}^m \in \arg \max_{\mathcal{L}' \in \Pi} \eta^{\mathcal{L}'}$, $m \geq M$.

根据式(4), 有

$$\begin{aligned} \omega_{e_{rs}^{kk'}}^m &= \frac{M}{m} \omega_{e_{rs}^{kk'}}^M + \frac{1}{m} \sum_{v=M}^{m-1} \hat{\omega}_{e_{rs}^{kk'}}^{v+1} = \\ &\frac{M}{m} \omega_{e_{rs}^{kk'}}^M + \frac{m-M}{m} \omega_{e_{rs}^{kk'}}^{\mathcal{L}} + \frac{1}{m} \sum_{v=M}^{m-1} \zeta_v, \end{aligned} \quad (10)$$

式中 $\zeta_v = \hat{\omega}_{e_{rs}^{kk'}}^{v+1} - \omega_{e_{rs}^{kk'}}^{\mathcal{L}}$. 由于 $\zeta_v (v = M, M+1, \dots)$ 是一个鞅差序列, 即 $\frac{1}{m} \sum_{v=M}^{m-1} \zeta_v \rightarrow 0, m \rightarrow \infty$, 由式(10)可知, 当 $m \rightarrow \infty$, 几乎确定 $\omega_{e_{rs}^{kk'}}^m \rightarrow \omega_{e_{rs}^{kk'}}^{\mathcal{L}}$. 结合定理1可得

$$\lim_{m \rightarrow \infty} \mathcal{L}^m = \mathcal{L} \in \arg \max_{\mathcal{L}' \in \Pi} \eta^{\mathcal{L}'}, \text{w.p.1.}$$

证毕.

5 仿真结果(Simulation results)

通过数值仿真实验对算法1进行评估. 仿真参数选取如下: 设 SON 由5个节点组成, 按功能分为3类 ($j = 1, 2, 3$), 提供2种类型的业务服务 ($i = 1, 2$). i 类业务的服务持续时间满足均值为 $\frac{1}{\mu_i}$ 的指数分布, $\mu_1 = 0.4, \mu_2 = 0.5$, 服务收益率分别为 $1 (i = 1), 2 (i = 2)$. 可行服务组合方式有两种 ($k = 1, 2$), 具体如表1所示.

表1 仿真参数

Table 1 Simulation parameters

组合方式	业务链路	服务类型	服务容量
1	l_1	1	3
	l_2	2	2
2	l_1	1	2
	l_2	2	2
	l_3	2	1

为评估算法的收敛性和自适应性, 在不同的业务请求到达率 λ 下 (这里用不同的业务请求到达率来反映不同的应用环境), 运用算法1对 DSC 策略进行优化. 算法仅依据对样本轨道的观察数据进行策略的改进, 在计算过程中没有提供诸如业务请求到达率、服务速率、事件到达率或状态转移速率等相关系统或环境参数的信息. 在仿真实验中, 算法1能够在有限迭代次数 (实验中通常 $M < 60$) 内收敛到最

优策略, 表现出良好的自适应性. 图1所示为算法1分别在 $\lambda = 2.3, 2.7$ 和 $3.3 (\lambda = \lambda_1 + \lambda_2)$ 时的执行情况.

通过比较实验评估算法1有效性. 图2所示为不同的业务请求到达率 λ 下, 由算法1产生的最优 DSC 策略和完全分享法 (CS)、固定分配法 (CP)^[13] 的应用效果. 通过比较显示运用算法1优化的 DSC 策略能够有效提高系统资源利用率. 这是由于 CS 没有考虑各业务之间的区分服务, CP 难以适应业务需求的变化, 而 DSC 策略既考虑了业务之间的区分服务, 又对业务需求的变化具有良好的自适应性.

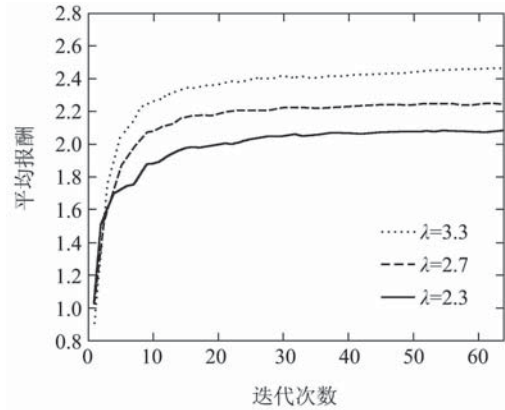


图1 算法1在不同应用环境中的执行情况

Fig. 1 Trajectories of average rewards when Algorithm 1 performed under different access arrival rates

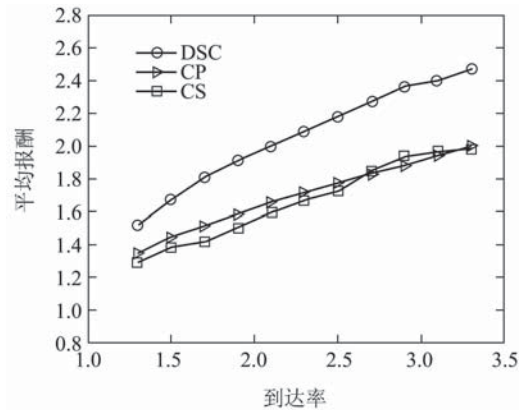


图2 应用效果比较

Fig. 2 Comparison of effects on system reward under various access arrival rates

6 结论(Conclusion)

本文针对复杂应用环境中网络新媒体服务系统的特点, 构建层次化的随机切换系统分析模型, 提出一种事件驱动的 DSC 策略在线策略迭代优化算法, 在提供各类业务 QoS 保证的同时, 有效提高系统资源的利用率. 该算法不依赖环境和系统参数的信息, 具有良好的自适应性. 采用事件驱动的策略有效缓解了网络化系统大状态空间导致的维数灾问题, 提高了算法的计算实时性, 有利于算法的在线实现.

在此模型的基础上可进一步考虑以各种统计类的QoS指标作为约束的DSC问题, 如不同类型业务的平均阻塞率要求小于相应的给定值. 优化目标即在满足统计类的QoS指标约束下, 使得系统的收益最大. 此时的最优策略属于随机型策略, 需要采用基于策略梯度的优化算法. 若考虑节省计算量, 可以采用策略迭代算法得到一个近似最优解.

参考文献(References):

- [1] SYCARA K, PAOLUCCI M, ANKOLEKAR A, et al. Automated discovery, interaction and composition of semantic web services[J]. *Journal of Web Semantics*, 2003, 1(1): 27 – 46.
- [2] SIRIN E, HENDLER J, PARSIA B. Semi-automatic composition of web services using semantic descriptions[C] // *Proceedings of the 1st Workshop on Web Services: Modeling, Architecture and Infrastructure*. Angers: ICEIS Press, 2003: 17 – 24.
- [3] SYCARA K, PAOLUCCI M, SOUDRY J, et al. Dynamic discovery and coordination of agent-based semantic web services[J]. *IEEE Internet Computing*, 2004, 8(3): 66 – 73.
- [4] SIRIN E, PARSIA B, WU D, et al. HTN planning for web service composition using SHOP2[J]. *Journal of Web Semantics*, 2004, 1(4): 377 – 396.
- [5] MCILRAITH S, SON T C. Adapting golog for composition of semantic web services[C] // *Proceedings of the 8th International Conference on Principles of Knowledge Representation and Reasoning*. Menlo Park: AAAI Press, 2002: 482 – 493.
- [6] 刘书雷, 刘云翔, 张帆, 等. 一种服务聚合中QoS全局最优服务动态选择算法[J]. *软件学报*, 2007, 18(3): 646 – 656.
(LIU Shulei, LIU Yunxiang, ZHANG Fan, et al. A dynamic web services selection algorithm with QoS global optimal in web services composition[J]. *Journal of Software*, 2007, 18(3): 646 – 656.)
- [7] DOSHI P, GOODWIN R, AKKIRAJU R, et al. Dynamic workflow composition using Markov decision processes[J]. *International Journal of Web Services Research*, 2005, 2(1): 1 – 17.
- [8] ZHAO H, DOSHI P. Composing nested web processes using hierarchical semi-Markov decision processes[C] // *Proceedings of American Association for Artificial Intelligence Workshop*. Menlo Park: AAAI Press, 2006: 75 – 83.
- [9] 范小芹, 蒋昌俊, 王俊丽, 等. 随机QoS感知的可靠web服务组合[J]. *软件学报*, 2009, 20(3): 546 – 556.
(FAN Xiaoqin, JIANG Changjun, WANG Junli, et al. Random-QoS-aware reliable web service composition[J]. *Journal of Software*, 2009, 20(3): 546 – 556.)
- [10] GU X, NAHRSTEDT K, CHANG R N, et al. QoS-assured service composition in managed service overlay networks[C] // *Proceedings of the 23rd IEEE International Conference on Distributed Computing Systems*. Piscataway: IEEE Computer Society, 2003: 1 – 8.
- [11] LIANG J, NAHRSTEDT K. Service composition for advanced multimedia applications[C] // *Proceedings of Conference on Multimedia Computing and Networking 2005*. Bellingham: Society of Photo-Optical Instrumentation Engineers, 2005: 228 – 240.
- [12] GU X, NAHRSTEDT K. Distributed multimedia service composition with statistical QoS assurances[J]. *IEEE Transactions on Multimedia*, 2006, 8(1): 141 – 151.
- [13] BANNAZADEH H, LEON-GARCIA A. Allocating service to applications using Markov decision processes[C] // *Proceedings of IEEE International Conference on Service-Oriented Computing and Applications*. Piscataway: IEEE Computer Society, 2007: 141 – 146.
- [14] CAO X R. Basic ideas for event-based optimization of Markov systems[J]. *Discrete Event Dynamic Systems: Theory and Application*, 2005, 15(6): 169 – 197.
- [15] GUO L, TAN E, CHEN S, et al. Does Internet media traffic really follow Zipf-like distribution?[C] // *Proceedings of the 2007 International Conference on Measurement and Modeling of Computer Systems*. New York: ACM Press, 2007: 359 – 360.
- [16] JIANG Q, XI H S, YIN B Q. Dynamic file grouping for load balancing in streaming media clustered server systems[J]. *International Journal of Control, Automation and Systems*, 2009, 7(4): 630 – 637.
- [17] CAO X R. Semi-Markov decision problem and performance sensitivity analysis[J]. *IEEE Transactions on Automatic Control*, 2003, 48(5): 758 – 768.
- [18] FANG H T, CAO X R. Potential-based online policy iteration algorithms for Markov decision processes[J]. *IEEE Transactions on Automatic Control*, 2004, 49(4): 493 – 505.
- [19] DAI G P, YIN B Q, LI Y J, et al. Performance optimization algorithms based on potentials for semi-Markov control processes[J]. *International Journal of Control*, 2005, 78(7): 801 – 812.

作者简介:

江琦 (1967—), 男, 副研究员, 主要研究方向为网络化系统的控制与优化, E-mail: jiangqi@ustc.edu.cn;

奚宏生 (1950—), 男, 教授, 博士生导师, 主要研究方向为离散事件动态系统、网络化系统的控制与优化, E-mail: xih@ustc.edu.cn;

殷保群 (1962—), 男, 教授, 博士生导师, 主要研究方向为离散事件动态系统、Markov决策过程, E-mail: bqyin@ustc.edu.cn.