

基于自律计算的系统服务可信性自优化方法

朱丽娜, 吴庆涛, 娄 颖, 郑瑞娟

(河南科技大学 电子信息工程学院, 河南 洛阳 471023)

摘 要: 提出了一种基于 Q 学习的系统服务性能自优化方法. 该方法通过感知网络系统的服务性能状态参数变化, 利用前馈神经网络的非线性映射关系得到执行动作, 综合系统服务性能变化情况及服务的可用性来计算环境奖赏函数值, 利用 Q 学习的自学习特性和预测能力, 使系统服务性能达到最优. 仿真结果表明, 该方法对优化系统整体可信性和服务效用具有明显的优越性.

关键词: 自律计算; 服务可信性; 前馈神经网络; Q 学习算法

中图分类号: TP393.08

文献标识码: A

文章编号: 1000-7180(2013)08-0063-04

Self-Optimization Method of System Service Dependability Based on Autonomic Computing

ZHU Li-na, WU Qing-tao, LOU Ying, ZHENG Rui-juan

(Electronic & Information Engineering College, Henan University of Science & Technology, Luoyang 471023, China)

Abstract: According to the notion of autonomic computing, a network system service performance online real-time optimization method is proposed based on Q-learning algorithm. First, service performances are taken as parameters, which can affect network system as targets. Second, feed forward neural network is used as nonlinear mapping to gain executive action. Finally, the environment reward function values are calculated according to the change of system service performance and service availability. Then self learning feature and predictive ability of Q-learning is used to make the system service performance achieve optimization. The simulation results show that the optimization method has obvious advantages in credibility and service utility of the overall system maintenance.

Key words: autonomic computing; service dependability; feed-forward neural network; Q-learning algorithm

1 引言

目前, 围绕自律计算技术和可信性相结合的研究被诸多研究机构和学者所关注. 文献[1]中探讨利用自律计算技术消除、避免系统的错误从而达到可信的目的. 文献[2]中通过设置自律管理框架使系统具有自动性能优化功能. 文献[3]中对自律可信性模型方面进行了研究.

采用“以技术管理技术”的主动式管理方法, 构造能够在弱化系统复杂性的同时, 为用户提供可信服务的系统, 主要面临以下问题: (1) 当所需服务可信性能下降时, 如何动态监视影响系统服务性能变

化的不确定系统状态参数, 实现及时感知; (2) 如何构建服务可信性能下降时的适应性调整过程, 以期寻找策略支持系统服务长期有效地可信运行. 为此, 本文从自律计算的角度, 提出一种基于 Q 学习的系统服务性能实时自优化方法, 从系统服务状态参数感知、“状态—动作”对选取、环境奖赏值计算以及 Q 学习算法的自主执行等方面, 实现系统服务可信性的自主优化.

2 基于 Q-learning 的系统服务可信性能自优化模型

Q 学习算法^[4]是一种与模型无关的强化学习算

法,它不需要估计环境模型,而是直接优化学习状态—动作对值,被认为是目前比较有效的一种强化学习算法.本文采用 Q 学习算法来调整影响系统的可信性能参数,并基于 Q 学习算法构建一个面向网络系统的服务性能实时自优化模型,其结构如图 1 所示.

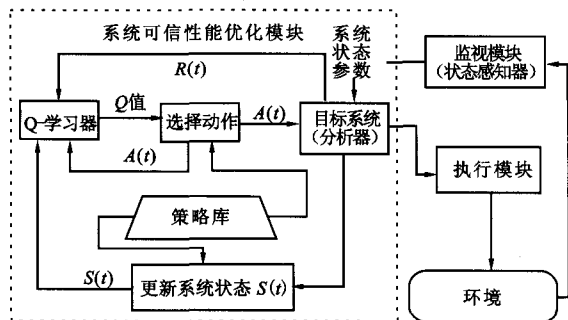


图 1 基于 Q-learning 的系统服务可信性能自优化模型

系统主要由以下三个模块组成:

(1) 监视模块: 监视模块通过部署在系统各个部分的状态感知器感知系统环境变化并采集影响当前系统服务性能的状态特征信息, 经过规范化处理后提交给目标系统的分析器。

网络系统服务的运行状况可用一组体现网络安全状态的观测变量表示, 设 t 时刻的状态观测序列为 S_t , 则集合 $S_t = \{s_1, s_2, \dots, s_m\}$, m 为系统提供的服务数目, S_i (其中 $i=1, 2, \dots, m$) 表示第 i 个服务的运行状态观测值, 且 $s_i = \{v_i | v_i \in \{v_1, v_2, \dots, v_n\}\}$, 其中 $\{v_1, v_2, \dots, v_n\}$ 为某特定服务的运行状态观测值集合。

(2) 自优化模块: 自优化模块是整个优化机制的核心, 其中的分析器对收集到的信息进行约减、聚类或融合处理, 筛选出影响系统服务性能下降的状态特征参数, 针对筛选出的状态特征参数, 依据存储在策略库中的可信策略, 利用前馈神经网络实现系统“状态—动作”对的非线性映射关系得到提高系统服务性能需要执行的动作, 并根据服务可用性和系统性能变化计算得到的环境奖赏函数值, 系统自主执行学习算法更新 Q 值。

(3) 执行模块: 将最优决策结果作用于系统, 完成对可信性的保持和增长。

以上三个模块共同构成一个自律反馈控制结构, 以终端网络系统为保护对象, 由自优化模块进行最优策略的选择, 即通过每隔一定决策周期根据当前系统性能状况自适应地对系统服务能力进行优化, 在实现可信性的同时减少了人为的干预。

3 网络系统服务性能实时优化算法

自优化模块根据环境状态特征参数到动作的映射, 采用“胜者为王”的方式确定具有最大概率值的执行动作, 通过对服务性能和服务可用性的改进获得长期的、最大的正的环境奖赏值, 执行学习算法更新值函数 Q, 获取最优策略并作用于系统, 达到增强系统可信性的目的。

3.1 “状态—动作”对选取方法

“状态—动作”对选取的基本思路是: 通常一个三层的前馈神经网络^[5]即可实现任意非线性映射, 利用前馈神经网络的三层自组织结构, 可以表示从状态感知器监视到的系统状态 S_t 到动作 $A(t)$ 的概率分布映射. 假定第一层为输入层, 该层神经元数为状态感知器所感知的系统状态数目加上动作变量的和, 输入向量为 M 维, 即 $x = (s_1, s_2, \dots, s_t, a_1, a_2, \dots, a_t)$. 隐层神经元数用 K 表示, 其大小可由文献^[6]中给出的经验公式(1)来确定:

$$K = 0.51 + \sqrt{0.43MZ + 0.12Z^2 + 2.54M + 0.77Z + 0.35} \quad (1)$$

式(1)中计算值需经四舍五入取整. 其中, M 为输入层神经元数, Z 为输出层神经元数(即动作个数). 用式(1)可确定隐层神经元数目的初值, 并可根据情况进行微调。

第 i 个隐神经元的输出值形式如式(2):

$$\rho_i = \vartheta(\bar{\omega}_{ij}\chi - \zeta_i^i) = \vartheta\left(\sum_{j=1}^N \bar{\omega}_{ij}\chi_j - \zeta_i^i\right), i=1, 2, \dots, K \quad (2)$$

式中, ϑ 为激励函数, $\bar{\omega}_{ij}$ 为第 i 个隐神经元与第 j 个输入神经元的连接权值, ζ_i^i 为该隐元的阈值. 则第 k 个输出神经元的输出值为

$$\Delta_k(x) = \mathcal{F}(\omega_k \rho - \zeta_k^z) \quad (3)$$

式中, $\mathcal{F}$ 为激励函数, $\rho = (\rho_1, \rho_2, \dots, \rho_N)$ 为来自隐层的输入向量, ω_k 为连接权值, ζ_k^z 为输出节点的阈值, 神经网络的输出 $\Delta_k(x)$, $k=1, 2, \dots, Z$ 表示动作 k 的概率值. 选取具有最大概率值 $\Delta_k(x)$ 的动作。

3.2 基于系统可信性能参数约束的评价函数

网络服务主要强调服务的可用性, 在开放网络中, 要保证目标系统在提供服务时的可信度需要考虑的参数很多, 通过系统运行时对服务可用性和网络服务性能参数的约束, 使反馈的评价函数所包含的网络状态信息更加全面. 为此, 根据影响目标系统可信性能参数约束的需要, 定义环境奖赏 R_t 为多数

据源参数性能约束的评价函数以满足系统服务可信性需求,假定 $R = \{R_1, R_2, \dots, R_t\}$ 表示每次执行动作后所对应的 R_t 集合。

本文以执行某个动作后获得长期的、最大的环境奖赏用于保证系统的可信性需求,若执行某个动作后获得环境正的奖赏,则有增强系统可信性的趋势,反之则降低这种趋势。环境奖赏函数值如式(4)所示:

$$R_t = \sum_{j=1}^m \omega(t_j) A_{jt} \cdot \Delta E_s \quad (4)$$

式中, $\omega(t_j)$ 为采取第 j 次行为/策略所占目标系统的比重, A_{jt} 表示系统服务的可用性, ΔE_s 值反映系统服务的性能变化。决定服务可用性的因素主要包括服务器 CPU、状态连接、内存占用率和负载等服务计算资源,因此,系统服务可用性计算可用式(5)表征:

$$A_s = 1 - \prod_{i=1}^m (1 - \omega_i \cdot A_i) \quad (5)$$

取 m 代表决定系统服务可用性的计算资源个数,一般取 4, A_i 为第 i 类计算资源的可用性情况, ω_i 表示各计算资源在服务可用性中的权重值。

通常,反映网络系统服务性能的参数主要有网络服务响应时间、传输速率均值和平均时延等。其计算公式如(6)所示:

$$\Delta E_s = \sum_{i=1}^n \omega_i \cdot \Delta E_i \quad (6)$$

式中, n 为反映网络服务性能变化的参数个数,通常取 3, ω_i 为该因素在服务性能中所占的权重, ΔE_i 为计量第 i 个因素在系统执行动作策略前后的性能变化差值。

3.3 算法描述

在确定了“状态—动作”对选取方法和奖赏函数值后,即可进行基于 Q 学习控制算法的动态实时优化。算法基本思想:设置 $Q_t(s_t, a_t)$ 的初始估计值 $Q_{t+1}(s_{t+1}, a_{t+1})$ 为 0,重复观察每个时刻当前环境状态 s_t ,选择并执行动作 a_t ,并根据计算得到的环境奖赏函数值 R_t ,系统自主执行学习算法更新估计值 $Q_{t+1}(s_{t+1}, a_{t+1})$,并更新系统状态 $s_{t+1} \leftarrow s_t$ 。其中值函数 Q 的更新公式为

$$Q_{t+1}(s_{t+1}, a_{t+1}) = (1 - \alpha) Q_t(s_t, a_t) + \alpha [R_t + \gamma \max_{A} Q_t(s_t, a_t)] \quad (7)$$

式中, $\gamma \in (0, 1)$ 是时间折扣因子, $\alpha \in (0, 1)$ 是学习速率, a_{t+1} 和 s_{t+1} 分别是下一时刻迭代动作和状态。其中 α 越大则收敛越快,但并不代表 α 越大越好,过大的 α 可能造成不收敛。Watkins 证明了 α 在满足

一定条件时,经过无数次迭代学习后,每个状态上最大的值函数会收敛到定值^[7]。

$$\max_{a \in A} Q(s_t, a_t) \rightarrow E \left\{ \sum_{n=1}^{\infty} \gamma^{n-1} R_{n+1} \right\} \quad (8)$$

其算法执行步骤如下:

步骤一:输入状态集合 $S_t = \{s_1, s_2, \dots, s_m\}_t$ 、动作集合 $A_t = \{a_1, a_2, \dots, a_t\}$,并初始化 $Q_{t+1}(s_{t+1}, a_{t+1}) \leftarrow 0$ 。

步骤二:由动作概率分布法式(3)选择并执行调度动作 a_t ;

步骤三:由式(4)得到一个奖赏函数值 R_t ;

步骤四:根据式(3)和式(4)更新 Q 函数值式(7);

重复执行以上步骤,当 $Q_t(s_t, a_t)$ 值达到最大时跳出执行第五步;

步骤五:确定 $\max_{a \in A} Q_{t+1}(s_{t+1}, a_{t+1})$,更新 $s_{t+1} \leftarrow s_t$ 。

4 仿真实验和结果分析

本文利用 MATLAB7.0 作为仿真实验工具,分别对优化算法前后,网络运行过程中对系统总体可信性的维护效果和系统服务效用维护效果进行仿真并进行对比。

系统的组成如图 2 所示,客户机通过局域网向服务器发出服务请求,两次对比实验分别针对传统服务器和增设了优化模型的服务器开展,仿真实验中,2 000 个服务随机地部署到局域网中构成服务资源,根据文献[8]的观察设置每天 2.3% 的服务失效率。

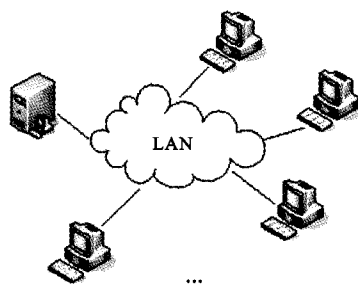


图 2 实验环境

第 1 组仿真实验的目的是为了验证应用自优化模型对系统整体可信性能的效果,图 3 所示为系统运行 90 天的系统服务性能平均可信性,图 3(a)为传统服务器对系统提供服务时可信性能的维护效果,图 3(b)为添加自优化模型后系统提供服务时可信性能的维护效果。

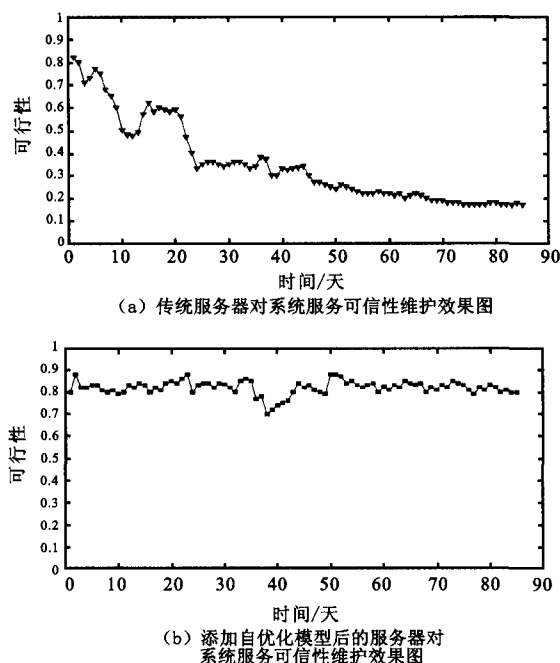


图3 效果图

通过图 3(a)和图 3(b)对比发现,传统服务器可信性能下降比较明显,添加自优化模型的服务器能够保持较好的可信性,波动范围较小,且趋向稳定在较高的水平。

第 2 组仿真实验分析了添加自优化模型前后对系统服务效用的影响。如图 4 所示,2 000 个服务构成的局域网服务资源以及 10 000 个请求,运行 90 天的平均效用。实验结果表明传统服务器的服务效用随可信度降低而明显下降,添加优化模型后的系统采用自优化控制方法,优化系统环境后,系统所提供服务能够保持较好的可信性,因而能够获得比较稳定的效用。

5 结束语

由于计算系统管理复杂性的急剧增加,导致传统使用复杂结构实现可信的方式难于满足现实的需要。本文提出了一种基于 Q 学习的网络系统服务可信性能的自优化方法,通过周期性对目标系统相关特征进行采集,并根据“状态—动作”对以及环境奖赏函数值,计算逼近最优函数值 Q,对其服务所处环境状态实时优化从而使服务可信性保持非降状态。

实验结果表明,具有优化模型的系统的总体可信性和服务效用明显优于未设置自优化模型的系统。在将来的研究工作中,将重点关注对关键状态参数指标的提取上。

参考文献:

- [1] Banks J, Crow E. Embedded diagnostics enable military ground vehicle autonomic logistics[C]//Proc of Reliability and Maintainability Symp (RAMS' 07). Piscataway, NJ: IEEE Society, 2007:48-52.
- [2] Hinchey M. Modeling for NASA autonomous nano-technology swarm missions and model-driven autonomic computing[C]//Proc of the 21st Int Conf on Advanced Information Networking and Applications (AINA'07). Piscataway, NJ: IEEE Society, 2007: 250-257.
- [3] Perter S, Narasimhan P, Wilkes J, et al. Prato: databases on demand[C]//Proc of the 4th Int Conf on Autonomic Computing (ICAC' 07). Piscataway, NJ: IEEE Society, 2007:11-15.
- [4] 余涛,甄卫国. 基于多步回溯 $Q(\lambda)$ 的 PSS 最优控制方法的研究[J]. 电力系统保护与控制, 2011, 39(3): 18-23.
- [5] 刘海涛,洪炳熔,朴松昊,等. 不确定环境下基于进化算法的强化学习[J]. 电子学报, 2006, 34(7): 1356-1360.
- [6] 徐明亮,须文波. 自适应 RBF 网络 Q 学习控制[J]. 控制与决策, 2010, 25(2): 303-306.
- [7] 洪渊,刘簪. 基于 Q 学习的虚拟人自适应感知系统建模[J]. 计算机应用, 2011, 31(1): 78-81.
- [8] 郭慧鹏,怀进鹏,邓婷,等. 一种可信的自适应服务组合机制[J]. 计算机学报, 2008, 31(8): 1434-1443.

作者简介:

朱丽娜 女, (1988—), 硕士研究生. 研究方向为网络信息安全.

吴庆涛 男, (1975—), 博士, 副教授, 硕士生导师. 研究方向为网络信息安全.

姜颖 男, (1980—), 博士, 讲师. 研究方向为数据管理, 信息检索.

郑瑞娟 女, (1980—), 博士, 副教授. 研究方向为网络信息安全.