

doi:10.3969/j.issn.1673-9833.2012.05.018

基于免疫代理的网络故障诊断研究

李岩岩, 文志诚, 吴欣欣, 叶健健

(湖南工业大学 计算机与通信学院, 湖南 株洲 412007)

摘要: 针对现有网络故障诊断系统的自适应能力弱、诊断速度慢和故障模式不同等问题, 将免疫原理与代理技术相结合, 采用分层多步的诊断思想构建诊断模型。基于克隆选择学说, 提出新算法, 完成检测器的训练。该算法选取了检测器克隆群, 引进优化参数, 可避免过早收敛和局部最优解的产生, 从而改进诊断性能; 同时还引入了检测器分类思想, 加速诊断过程。与传统故障诊断方法相对比, 该模型在处理复杂环境下的网络故障具有较明显的优势。

关键词: 网络故障诊断; 免疫原理; 代理技术

中图分类号: TP393

文献标志码: A

文章编号: 1673-9833(2012)05-0081-06

Research of Network Fault Diagnosis Based on Immune Agents

Li Yanyan, Wen Zhicheng, Wu Xinxin, Ye Jianjian

(School of Computer and Communication, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: Aiming at the weak adaptive capacity, slow diagnosis and different fault modes for the existing network fault diagnosis system, combines the immune principle and agent technology and uses hierarchical multiple step diagnosis concept to construct diagnosis model. Based on the clone selection theory, proposes a new algorithm to complete detector training. The algorithm selects the detector colony and introduces dynamic optimizing parameter, which avoiding premature convergence and generation of local optimal solution, greatly improves the diagnostic performance. The introduction of detector classification concept accelerates the diagnosis process. Comparing to the traditional fault diagnosis methods, proves that the model in the treatment of complex environment network fault has obvious advantages.

Keywords: network fault diagnosis; immune principle; agent technology

0 引言

现在已有许多智能诊断方法运用于网络故障诊断, 如基于模型的方法、专家系统、神经网络技术、遗传算法等^[1]。这些方法虽然为精确诊断设备的故障起到了重要作用, 却存在着一些不足, 如系统自学习和自适应能力较弱, 缺乏并发诊断能力, 难以

适应动态变化的环境和求解复杂系统的全部任务要求, 对异常类故障的诊断依然面临困难, 因此, 寻求高效智能的故障诊断方法成为研究热点。计算机的故障诊断与生物免疫系统^[2]所遇到的问题具有惊人的相似性, 两者都要在不断变化的环境中维持系统的稳定性。Agent 技术^[3-4]具有分布性、灵活性、自适应性和鲁棒性等特点, 正是网络故障诊断领域

收稿日期: 2012-05-12

作者简介: 李岩岩 (1986-), 女, 辽宁铁法人, 湖南工业大学硕士生, 主要研究方向为软件工程和可信软件,

E-mail: 1106018836@qq.com

所期望得到的。

生物信息系统可看成一分布式自治系统，并且能提供给科学家和工程领域各种富有成效的技术和方法。因此，Agent技术与生物免疫系统信息处理机制相结合为智能故障诊断系统的设计和实现提供了一条极具潜力的途径。本文在深入分析生物免疫机理和研究具有高度自适应、自控制、自学习的Agent技术的基础之上，针对目前网络拓扑结构和网络故障的特点，构建了一种免疫算法与Agent技术相结合的网络故障诊断模型。该模型针对复杂多变的网络环境，采用分层逐步的诊断思想进行由浅入深地逐层诊断，可提高网络故障诊断的确诊率。

1 网络故障诊断系统模型

1.1 系统模型介绍

根据系统故障诊断和生物免疫系统的特点，以Agent的基本结构作为框架，采用逐层分步的诊断思想，并在每个层次步骤中采用Agent联盟的诊断策略，完成对故障的诊断。诊断模型可看作是由单元组形成的多Agent联盟诊断模型，具体过程如下：当网络出现故障时，其信息经过处理后，作为抗原被提交给本地Agent，本地代理与本地知识库进行匹配，若匹配成功，则给出故障类型，否则将全部信息提交给诊断中心；诊断中心根据诊断任务和本地资源，确定实施该任务的代理集合，并进行相应的角色分配，这些携带需要执行诊断任务的移动代理抵达目标节点后，再和当地的SNMP（simple network management protocol）代理交互，获取需要的实时网络状态信息，再由数据采集系统代理对故障信息执行存储及预处理操作；各诊断Agent根据获得的实时信息分别利用各自不同的知识库和推理机制对故障状况进行并行推理、独立求解，将多个Agent的结果（故障的分类信息）作为多源证据，通过结果融合^[5]，得出比任意Agent更高精度的诊断结论；若无法给出诊断结果，就将保存在代理中的全部数据和节点响应消息带回到诊断中心，返回信息作为抗原，经过诊断中心故障诊断模型处理后，根据不同的诊断状况，可以得到3种结果：无故障，发生已知故障，发生未知故障。

1.2 系统流程设计

免疫代理的网络故障诊断系统的工作流程见图1。该系统主要包含2功能：1) 派遣到节点环境诊断代理集合；2) 诊断中心代理集合。

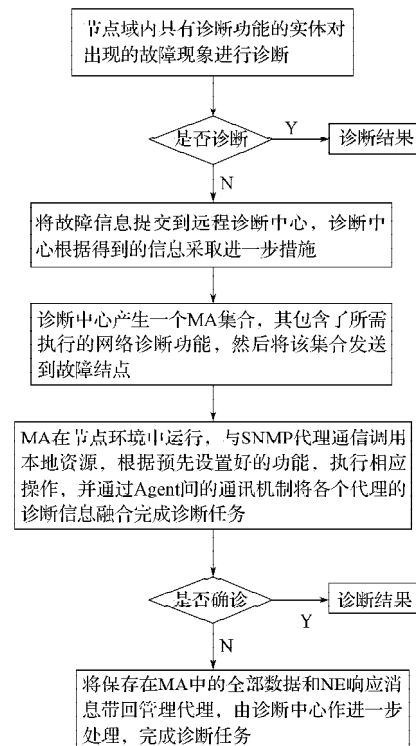


图1 系统流程图

Fig. 1 Flow chart of system

1.2.1 派遣到节点环境诊断代理集合

代理集合结构见图2。

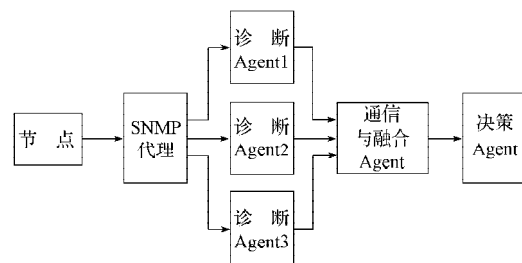


图2 派遣到节点环境诊断代理集合图

Fig. 2 The collection of diagnostic agents sent to the node environment

该诊断代理集合具有以下优点：

- 1) 利用与本地SNMP代理交互获得的数据相对来说是比较新的，能够更好地反映当前网络的现状，实时性方面较以往得以提高。
- 2) 可以依据出现的故障特征组建立相应的诊断代理，因此诊断机制更加灵活、准确，而且易于修改、扩充。

1.2.2 诊断中心网络模型

诊断中心网络模型，采用逐层分步诊断联盟的思想，系统在不同智能层次的行为方式是不同的，因此，对于不同的智能层次可以结合不同的免疫算法采用不同的处理方式，从而构造出逐层诊断联盟的Agent诊断结构，此诊断结构主要由3个代理完成对故障信息的深层诊断。

1) 第一个代理。该代理相当于生物免疫系统的固有免疫,主要是检测已知的曾经发生过的故障,这个代理的抗体库中包含自体检测器和故障检测器,便于快速检测出已知的故障。

2) 第二个代理。该代理作用是避免错误肯定与错误否定的现象发生。在实际中,由于计算机网络环境是一个高度复杂且动态变化的系统,自体集和非自体集不可能是静态不变的,对于无法确定的故障信息,根据其特征,选择合适的模糊判断法,对其推理诊断,若诊断结果为正常行为,则给予警告,否则确认为新的故障现象出现。

3) 第三个代理。该代理主要功能是以自适应的方式解决第一层、第二层未能识别的新故障类型。通过算法设计对未知故障进行学习、推理,并产生相应抗体,完成对未知故障的识别诊断。

诊断中心网络模型如图3所示。

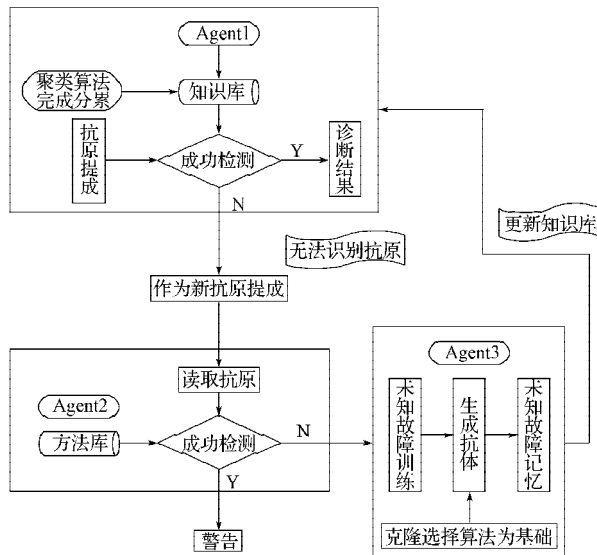


图3 诊断中心网络模型

Fig.3 Network model for diagnostic center

2 检测器的生成算法

2.1 算法的引入

未知故障检测器的生成对未知故障的确定具有至关重要的作用。在研究现有检测器生成算法的基础上,针对现有算法容易陷入局部极值点等缺陷,本文根据生物免疫系统中的学习机理及其相关的免疫学理论中的克隆选择学说^[6-7],提出了一种新的用于未知故障检测器生成的免疫算法。该算法主要包括选择、克隆扩增、检测器的优化等。根据算法亲和度自适应调节学习参数的方法,以提高算法的全局寻优能力。通过算法中的克隆和变异操作,将检测器群进行局部扩展;通过选择删除操作,将扩展后

的检测器种群进行压缩,获得优秀的候选解集;通过反复的扩张与压缩,得到最优解。

1) 选择候选种群。选择候选种群是在抗体群 M 中选择一个抗体进入克隆抗体群 C 中的随机映射。抗体群 M 来自诊断中心 Agent1 的检测器集合,包括自体检测器和已知故障检测器。 n_{ij} 表示抗体 m_{ij} 对抗原 x 学习效果的学习参数,为二者亲和度函数,即

$$n_{ij} = f_{ij} / f_{ave}, \quad (1)$$

式中: f_{ij} 为抗体 m_{ij} 与抗原 x 的亲和度,即

$$f_{ij} = 1 / (1 + d_{ij}), \quad d_{ij} = \sqrt{\sum_{i=1}^m (m_{ij} - x)^2}, \quad 1 \leq i \leq m, 1 \leq j \leq m;$$

$$f_{ave} = (f_{i1} + f_{i2} + \dots + f_{im}) / m.$$

β 为候选种群能否作为抗体的阈值。当 $n_{ij} \geq \beta$, 抗体被选入克隆抗体群 C 。由上述过程可以看出,学习效果是在算法的执行过程中由抗体与抗原的亲亲和度体现出来。

2) 克隆检测器。将选取的 c 个检测器克隆,产生克隆集合 C' , 克隆的数目采用非等比例方法^[8],即克隆的数目由亲和度决定,亲和度越大的克隆数目越多。克隆集的选取方法为:根据亲和度的大小对检测器进行排序,将检测器群体按亲和度大小分为3个等级,抽取第一个等级前30%的高亲和力检测器,第二个等级前20%的高亲和力检测器,第三个等级前10%的高亲和力的检测器,将抽取的这3类检测器作为克隆集。克隆后得到的抗体总数为

$$c' = c_1 + c_2 + c_3, \quad (2)$$

式中: c' 表示克隆总数, c_1, c_2, c_3 分别表示检测器群体按亲和力大小形成3个等级的克隆数目,即

$$c_1 = \left\lfloor 0.3n_{m_1} \frac{f_{ij}}{\sum_{j=1}^{m_1} f_{ij}} \theta_{ij} \right\rfloor, \quad (3)$$

$$c_2 = \left\lfloor 0.2n_{m_2} \frac{f_{ij}}{\sum_{j=m_1}^{m_2} f_{ij}} \theta_{ij} \right\rfloor, \quad (4)$$

$$c_3 = \left\lfloor 0.1n_{m_3} \frac{f_{ij}}{\sum_{j=m_2}^{m_3} f_{ij}} \theta_{ij} \right\rfloor; \quad (5)$$

$\lfloor \bullet \rfloor$ 表示向下取整;

n_{m_i} 表示与克隆规模有关的设定值;

θ_{ij} 表示抗体 m_{ij} 与其他抗体间的亲和力,反应了

抗体之间的相似度, 即 $\theta_{ij} = \min_{j=1, \dots, m} \{d_{ij}\}$ 。

通过克隆操作, 形成了子抗体群 C' , 抗体总数为 c' 。

3) 优化检测器。最优抗体的产生主要以变异为主, 本算法引入动态变异算子, 在算法运行的初期, 提高变异算子的变异概率, 当算法运行到后期尤其是将收敛到近似解时降低变异算子的变异概率。引入动态变异算子的好处是一方面实现多目标优化, 跳出局部极值点, 另一方面提高算法的收敛速度。动态变异概率为

$$m'_{ij} = m_{ij} + pN(0,1), \quad (6)$$

式中: m'_{ij} 为变异后的检测器;

m_{ij} 为变异前的检测器;

$N(0, 1)$ 是服从均值为 0, 方差为 1 的正态分布随机变量;

p 为个体变异率,

$$p = p(1 - t \times \alpha / T); \quad (7)$$

t 为当前进化代数;

T 为总的进化代数;

α 为常数。

4) 检测器评价。计算变异后的检测器与抗原匹配程度, 如果变异后检测器与抗原的亲合程度低于变异前的, 则删除变异后的检测器, 并用随机生成的检测器加以替换, 从而保持检测器群体的多样性, 即

$$C = C \cup \{m_{ij}\}, \quad (8)$$

式中, m_{ij} 的选择取决于 f_{ij}^* , $f_{ij}^* = \max\{f_{i1}^*, f_{i2}^*, \dots, f_{ic'}^*\}$, 其中, c' 为克隆扩增后检测器群 C 的规模, “*” 为经过克隆扩增和检测器优化操作。

2.2 算法的具体实现步骤

诊断中心未知故障检测器的生成算法见图 4。

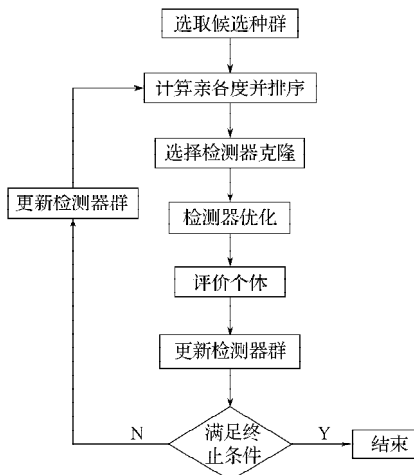


图 4 免疫算法流程图

Fig. 4 Flowchart of immune algorithm

免疫算法的具体步骤如下:

步骤 1 选取候选种群, 根据所要求解的问题, 在自体检测器和已知故障检测器选取 N 个体作为初始群体。

步骤 2 计算选取的检测器与抗原间的亲和度, 并按照亲和度大小排序、分三级。

步骤 3 取分级后按亲和力由高到低的顺序, 第一个等级的前 30%, 第二个等级的前 20%, 第三个等级的前 10% 作为克隆对象。

步骤 4 按照式 (6) 对克隆后的群体进行优化。

步骤 5 计算变异后的检测器与抗原的匹配程度, 按亲和力最大原则, 选择最佳抗体。

步骤 6 保留新产生的最佳抗体, 抑制清除其他所有抗体, 更新检测器群。

步骤 7 条件判断。如果满足抗体最优条件, 算法结束, 输出结果; 如果没有达到最优条件, 转至步骤 2 继续。

部分代码如下:

BEGIN

Select the candidate population ;

LOOP: Calculate affinity and sort;

Select detector clone;

Detector optimization;

Evaluation of individual;

Update the detector group;

IF (meet the antibody optimal conditions)

THEN End up algorithm, output the results;

ELSE GOTO LOOP;

END。

3 仿真实现与算法比对

3.1 仿真实现

该实验主要是实现诊断中心网络模型对网络故障的识别及存储。本文选择 IBM 公司开发的 Aglet 2.0.2 作为实验平台。实验定义了 7 种网络状态, 包括正常状态、接口不通、带宽不足、通信协议不兼容、缓冲区不足、接口性能问题、线路故障, 分别用 $C_1, C_2, C_3, C_4, C_5, C_6, C_7$ 表示。每个故障类型有 7 个属性 ($S_{1i}, S_{2i}, S_{3i}, S_{4i}, S_{5i}, S_{6i}, S_{7i}$), 依次是接口工作状态、输入丢弃率、输出丢弃率、输入错误率、输出错误率、网络利用率和未知协议率。接口的工作状态可以通过变量 ifAdminStatus, ifOperStatus 的综合取值获得, 2 个变量均包括 up, down, testing 3 种状态, 为方便处理, 2 个变量都处于 up 状态时视为正常, 接口工作状态的取

值为0.1,其他组合状态均视为故障状态,其取值为0.9。通过分析发现接口性能问题和线路故障表现出来的故障征兆完全相同,因此,将这2类故障作为一类考虑。故障诊断界面如图5所示。

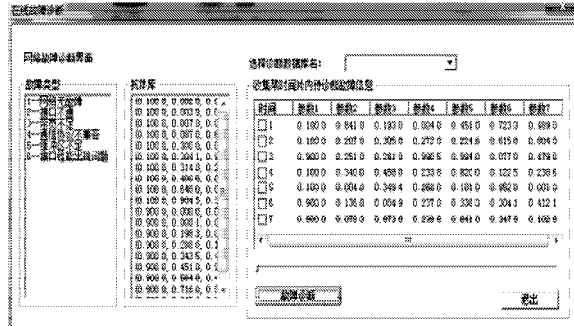


图5 故障诊断界面

Fig. 5 Interface of fault diagnosis

3组测试过程和结果如下:

1) 第1组测试。输入数据为(0.100 0, 0.200 0, 0.100 0, 0.300 0, 0.200 0, 0.100 0, 0.000 0),该测试结果网络无故障,如图6所示。



图6 网络无故障

Fig. 6 Network without fault

2) 第2组测试。输入数据为(0.100 0, 0.007 0, 0.006 0, 0.000 3, 0.005 0, 0.800 0, 0.000 0),该测试结果诊断为已知故障,如图7所示。



图7 诊断到已知故障

Fig. 7 Diagnosis to the known fault

3) 第3组测试。输入数据为(0.100 0, 0.904 5, 0.383 0, 0.402 0, 0.681 0, 0.106 0, 0.834 0),该测试结果诊断为未知故障,如图8所示。



图8 诊断到未知故障

Fig. 8 Diagnosis to the unknown failure

共测试了140组数据,其中未知故障不在故障类型中,测试结果见表1。

表1 测试结果

Table 1 Test results

故障类型	识别率/%
C ₁	95
C ₂	90
C ₃	90
C ₄	80
C ₅	90
C ₆	60
未知故障	80

由以上实验结果可知,基于免疫代理的网络故障诊断模型克服了传统网络故障诊断方法的缺陷,对于新发生的故障类型,也能够及时学习记忆,而且不再单纯依赖远程监测代理实现通讯,可以通过定义移动代理实现故障的监测、捕获和数据包处理,从而使其故障诊断更加灵活高效。该算法对网络接口组故障诊断的总体识别率较高,但对接口问题或网络线路问题故障识别率一般。

3.2 算法对比分析

针对确定的网络故障样本,在同样的硬件环境中,将本文提出的分层多步网络故障诊断模型(算法1)和传统的克隆选择算法(算法2)进行仿真比对。对2种算法分别测试了140组网络故障数据,7种故障类型,每种类型测试了20组数据,2种方法对网络故障诊断的正判情况见表2。

表2 诊断结果对比表

Table 2 The contrast of diagnostic results

故障类型	测试样本数	算法1		算法2	
		正确个数	正确率/%	正确个数	正确率/%
C ₁	20	19	95	18	90
C ₂	20	18	90	16	80
C ₃	20	18	90	18	80
C ₄	20	16	80	15	75
C ₅	20	18	90	16	80
C ₆	20	12	60	18	90
C ₇	20	16	80	14	70

从实验结果可以看出, 算法 1 具有较好的障诊断率, 但是算法 2 对网络的接口性能问题识别更好。

4 结语

本文设计的故障诊断系统在处理复杂环境下的网络故障时具有较明显的优势。它既借鉴了生物免疫的自学习、自适应和自组织的特性, 又利用了 Agent 的移动性、灵活性和快速性, 在故障诊断过程中, 实现了诊断的连续化、智能化及系统化, 而且通过算法设计, 能较准确快速地应对未知故障现象, 实现诊断系统全方位的智能诊断。

参考文献:

- [1] 张甜, 许家怡. 几种智能技术在网络故障诊断中的应用研究[C]//第十二届全国青年通信学术会议. 北京: 电子工业出版社, 2007: 1844-1848.
Zhang Tian, Xu Jiayi. Application Study of a Few Intelligent Techniques in the Network Fault Diagnosis[C]//The 12th National Youth Conference on Communication. Beijing: Publishing House of Electronics Industry, 2007: 1844-1848.
- [2] 李涛. 计算机免疫学[M]. 北京: 电子工业出版社, 2004: 19-20.
Li Tao. Computer Immunology[M]. Beijing: Publishing House of Electronics Industry, 2004: 19-20.
- [3] 何永勇, 褚福磊, 陈真勇. 基于多 Agent 的分布式故障智能诊断原型系统研究[J]. 计算机工程与科学, 2002, 24(1): 88-92.
He Yongyong, Chu Fulei, Chen Zhenyong. On Intelligent Diagnosis Prototype Systems for Multi-Agent Based Distributed Faults[J]. Computer Engineering & Science, 2002, 24(1): 88-92.
- [4] Wooldridge Michael. Intelligent Agents: The Key Concepts [J]. Multi-Agent Systems and Applications II, 2002, 2322: 3-43.
- [5] 钟联炯, 孙璐. 基于MAS的故障诊断任务分解和结果综合[J]. 西安工业大学报, 2007, 27(2): 176-180.
Zhong Lianjiong, Sun Lu. Multi-Agent System Based Task Decomposition Solution Synthesis in Fault Diagnosis[J]. Journal of Xi'an Technology University, 2007, 27(2): 176-180.
- [6] De Castro Leandro Nunes, Fernando J Von Zuben. The Clonal Selection Algorithm with Engineering Applications [C]//Workshop Proceedings of GECCO. Las Vegas: Morgan Kaufman Publishers, 2000: 36-37.
- [7] 陈文鑫, 陈军敢, 杨亚萍. 基于亲和力阈值的静态克隆选择算法[J]. 计算机工程, 2008, 34(10): 168-170.
Chen Wenxin, Chen Jungan, Yang Yaping. Static Clonal Selection Algorithm Based on Optimized Match Threshold [J]. Computer Engineering, 2008, 34(10): 168-170.
- [8] 陈婕, 熊盛武, 林婉如. NSGA-II算法的改进策略研究[J]. 计算机工程与应用, 2011, 47(19): 42-45.
Chen Jie, Xiong Shengwu, Lin Wanru. Improved Strategies and Researches of NSGA-II Algorithm[J]. Computer Engineering and Applications, 2011, 47(19): 42-45.

(责任编辑: 邓彬)