基于免疫算法的迁移工作流实例安全工作位置选择①

李 浩，韩芳溪，王晓琳

(山东大学 计算机科学与技术学院，济南 250061)

摘 要：在迁移工作流系统中迁移实例的安全和路线规划对于工作流目标的完成起着至关重要的作用。提出一种基于免疫算法的迁移策略，该方法借鉴生物免疫系统的特性，将迁移实例的需求以及安全要求编码为抗原，将工作位置的服务信息以及安全性能编码为抗体，模拟免疫系统的匹配过程从而动态的得到迁移路线。该方法不仅考虑迁移实例的任务需求，而且对迁移实例的需求信息划分安全等级，考虑工作位置的安全性，从而为迁移实例得到一个相对安全的迁移目标。仿真实验表明，该算法有效的避开了恶意主机，使迁移实例的被攻击率大大降低，保护了迁移实例的安全。

关键词：迁移工作流；迁移实例；免疫算法；迁移策略；安全

Migrating Workflow Instance's Selection of Security Workstation Based on Immune Algorithm

LI Hao, HAN Fang-Xi, WANG Xiao-Lin

(School of Computer Science and Technology, Shandong University, Jinan 250061, China)

Abstract: In migrating workflow system, the security and route of migrating instance for the completion of the workflow plays a vital role. This paper presents a migration strategy based on immune algorithm. This method draws on the immune system's features, encodes migrating instance's needs and security requirements as antigen and work location of service information and security as antibody, and simulates the matching process of the immune system in order to get the migration routes dynamically. The algorithm considers both the resource requirements of Agent's task and the security of the workstation according to Agent's risk in order to achieve safer route planning. The simulation results show that the algorithm effectively avoids the malicious host, significantly reduces the probability of migrating instance attacked and enhance the security of migrating instance.

Key words: migrating workflow; migrating instance; immune algorithm; migration strategy; security

1 引言

迁移工作流是近年来工作流管理系统研究的一个新方向。文献[1]给出了一个迁移工作流管理系统框架，该框架由一个迁移工作流管理机和若干个建立友好关系的局域网组成。在迁移工作流中，任务执行主体称作迁移实例(Migrating Instance)，迁移实例可以在某个工作位置上利用本地资源执行一项或多项任务。当迁移实例发现当前工作位置不能满足其执行任务的要求时，便携带任务说明书和当前执行结果迁移到另一个能满足其要求的工作位置上继续执行。由于迁移实例的工作行为在许多方面类似于移动 Agent（Mobile Agent，MA），因此迁移实例是以移动 Agent 为计算范型构造的，可以基于移动 Agent 的研究成果来研究迁移实例。MA 是一个能在异构网络中自主地从一台主机迁移到另一台主机，并可与其他 Agent 或资源交互的程序[2]。

移动性是 MA 最重要的特性之一[3,4]，是指 MA 可以携带自身状态和代码从一个环境转移到另外一个环境，并恢复执行。当 MA 发现当前主机无法满足服务需求时就会提出迁移申请，MA 迁移所经过的主机序

① 基金项目:山东省自然科学基金(ZR2009GM021)
收稿时间:2010-12-08;收到修改稿时间:2011-01-07

列就构成了一条迁移路径。迁移实例在迁移过程中，能自主的选择服务节点主机并移动到服务节点上完成所携带的任务。如何根据任务的需求以及动态变化的系统状况规划迁移实例本身的移动路线，是迁移工作流研究的重要核心问题之一。此外，由于迁移实例本身携带数据，容易遭到恶意主机的攻击(如信息窃取、数据篡改等攻击)，因此路径规划的另一个重点就是使迁移实例迁移到相对安全的主机上，这样才能够在一定程度上防患于未然。

K.Moizumi[5]等人根据当前网络负载和信息在主机上的分布概率为 Agent 规划了静态迁移路径，但没有考虑网络主机的动态性和 Agent 的迁移经验。刘大有、党辰和张正球[6-8]等人采用了不同迁移策略支持 Agent 动态迁移规划，却忽略了 Agent 在迁移过程中自身的安全问题，导致 Agent 可能遭受恶意主机的攻击。李鲁艳[9]提出基于任务片的旅行图生成算法，通过减少迁移次数来降低移动故障和移动安全导致的风险，但并不能保证 Agent 不会迁移到恶意主机上。

免疫算法是基于人体免疫系统中抗体与抗原识别、抗体产生过程等抽象出来的智能算法，在免疫算法解决问题时，抗原对应要解决问题的数据输入，如目标函数、约束等；抗体对应问题的解[10]。免疫算法的优良特性(如全局求解，自我调节和学习记忆等[11])使免疫算法在机器学习、生产调度、结构优化设计等领域问题求解中显出了优良的性能和效率[12,13]。免疫算法诸多特性能够更好的支持迁移实例的规划性能，工作流系统中工作位置众多，将工作位置信息编码为抗体，由迁移实例携带，一方面避免了迁移实例与服务引擎之间的过多交互，另一方面迁移实例可以动态地更新抗体信息，使抗体库具有自我调节和学习的能力，更能反映出当前状态下工作位置的真实情况。

本文提出一种基于免疫算法的迁移策略，该方法借鉴生物免疫系统的特性，将迁移实例的需求以及安全要求编码为抗原，将工作位置的服务信息以及安全性能编码为抗体，模拟免疫系统的匹配过程从而动态的得到迁移路线。该方法不仅考虑迁移实例的任务需求，而且对迁移实例的需求信息划分安全等级，考虑工作位置的安全性，从而为迁移实例得到一个相对安全的迁移目标。

2　问题描述

为了使迁移实例能够选择相对安全的主机或路径进行迁移，本文提出了一个基于免疫算法的安全迁移规划模型，在该模型中，迁移实例携带抗体库及免疫模块进行迁移，抗体库中包含有抗体，一个抗体对应一个工作位置，通过位置－服务关系构建抗体；免疫模块负责抗体的维护、亲和度的计算等等，通过对工作位置历史数据和当前状态的评价可以得到抗体信息，通过免疫模块可以得到当前状态下最佳的抗体。迁移实例提出迁移申请，会根据自身任务－服务关系构建抗原，并提交给免疫模块，由免疫模块通过免疫选择得到最佳的抗体并返回给迁移实例，迁移实例迁移后将迁移结果反馈给免疫模块，由免疫模块再对抗体库进行维护和更新。

定义 1. 抗原 Ag 对应迁移实例当前任务的服务需求以及安全需求，编码方式为(Id，SC(type, res, t)，[AL(A, L)])，其中 Id 为迁移实例的 Id；SC 为迁移实例的当前服务需求，表示为 SC(type, res, t)，type 是所需服务的类型，res 表示该类服务约束，包含该类型服务的所需参数，t 是迁移实例期望的运行时间；AL 为该迁移实例可能会遭受到的攻击，可能有多个，表示为 AL(A, L)，attack 代表攻击类型，L 为该攻击的危险级别，分为高、中、低三级(用攻击发生的频率来分级，分别是 0.05，0.1，0.2)，例如在查询图书价格服务中，该迁移实例可能遭受到修改信息、欺骗服务或窃取数据等攻击，但前两者的危险程度显然比后者高。

例如某抗原编码为(agent1, [orderbook, ("java", 25,600)], [(001,0.1),(002,0.05),(003,0.1)])，代表该抗原需要购书这项服务，其中服务内容为购买"java"这本书，该书价格下限为 25，购买数量为 600 本，该项服务可能会遭受 3 种类型的攻击，危险级别均较高。

定义 2. 抗体 Ab 对应工作位置提供的服务信息以及安全性能，编码方式为(SId，[Service(type, cap,)]，[Illegal(I, R)]，Type，Success，Lifespan)，其中 SId 为工作位置的标识；Service 为该工作位置所提供的服务(可能有多种服务)，对应于抗原中的 SC，包括服务类型 type、服务能力参数 cap，是指这个工作位置上该服务类型平均的运行时间；Illegal 为该工作位置的违规操作，包含该工作位置的违规类型 I(对应抗原中攻击类型)和 I 在最近时间 T 内发生的频率 R；Type 为该抗体类型，根据工作位置的执行次数、执行成功率和安

全系数可以将抗体类型分为未成熟抗体，成熟抗体，记忆抗体；Success 为该抗体所代表的主机的成功执行率；Lifespan 为该抗体的生命力，初始时为一固定值。

例如某抗体编码为(bookstore, [orderbook, ("java", 25,1000)], [(001,0.25),(002,0.02)，(003,0.005)，(004,0)], immature, 0.65, 10)，代表该抗体是一个书店，书店的服务成功率为 0.65，提供"orderbook"这项购书服务，其中"java"这本书的价格为 25，数量为 1000 本，该书店在过去的一段时间内共有 4 种违规类型，该抗体是不成熟的抗体，生命力为 10。

亲和度计算 亲和度是抗体满足抗原的程度，在本文中，抗体需在服务和安全两个方面满足抗原。由于编码使用实数表示，所以抗体和抗原亲和度以 Euclid 距离计算。根据上面的定义，亲和度分为服务和安全两个方面，其中服务亲和度的计算涉及到服务质量、通讯时延、服务成功率等等，安全亲和度为抗原的安全需求与抗体的违规操作之间的计算，亲和度的计算具体如公式(1)所示：

$$Aff(B)=\alpha(success\cdot\sum_{i=1}^{m}\frac{cap_i}{res_i}+\frac{t}{T})+\beta\sum_{j=1}^{n}\omega_j\frac{L_j-R_j}{L_j} \quad (1)$$

其中 α , β 和 ω_j 均为权重系数，且 $\alpha+\beta=1$， $\omega_1+\omega_2+...+\omega_n=1$， ω_j 的大小由抗原中该攻击类型的危险级别确定，危险越高，权重越大。定义 L_j 为抗原中攻击类型 j 所容忍的安全级别，对应的 R_j 为抗体中攻击类型 j 在 T 时间内发生的频率(j=1,2,3,...,n)；定义 cap_i 为抗体所代表的工作位置中服务类型 i 的服务能力，success 是该工作位置的执行成功率，res_i 为抗原中所对应的服务 i 的需求(i=1,2,3,...,m)；定义 t 是迁移实例期望的运行时间，是该工作位置上迁移实例平均的运行时间。抗原和抗体的亲和度越大，标明抗体所对应的工作位置越符合迁移实例的需求。

3 抗体演化

抗体的产生 抗体的产生有两种方式：1)针对待求解的问题，若能与已存在的抗体相匹配，则该抗体即为问题的一个解，寻找抗体次序为：记忆细胞、成熟细胞、未成熟细胞；2)否则，从服务引擎上随机选取一个能够完成该迁移请求的主机，并为该主机建立抗体信息，将该抗体放入未成熟细胞；

抗体的更新 当抗体与抗原成功的匹配并且迁移实例迁移成功以后，改变其执行成功率，并且按照下式更新其生命力：

$$lifespan\ (t+1)=lifespan\ (t)+\mu\cdot T \quad (2)$$

μ 为成功抗体的生命力调整系数，T 为执行的时间；而按照下式每隔一定的时间 Δt 衰减其生命力：

$$lifespan\ (t+\Delta t)=(1-\rho)lifespan\ (t) \quad (3)$$

ρ 为生命力衰减系数，ρ 增大，生命力衰减的快，ρ 减小生命力衰减的慢。所以为了保证抗体的数量的稳定，当抗体库中抗体的数量较少时，ρ 减小；当抗体库中抗体的数量较多时，ρ 增大；然后，更新抗体的类型以及环境信息，环境信息包括这段时间内该主机的服务信息和违规信息。

抗体的迁移 可以在相互信任的主机之间交换他们各自的优秀抗体，这样就使抗体库中的优秀抗体更多，但是迁移的次数不能过于频繁，另外迁移的范围也不能过小，否则迁移的效用不大，例如在两个联盟之间交换抗体就比联盟内交换抗体的效用更大。

抗体的生命周期 在抗体新创建时都会被赋予一个初始的生命值，随着时间的增加，抗体的生命力可能会减至 0，这就导致了抗体死亡，如图 1 所示，但是成熟抗体和记忆抗体中的优秀抗体可以作为新的抗体再加入到未成熟抗体中，这样就保证了在一段时间内没有被使用的抗体不会完全消亡。

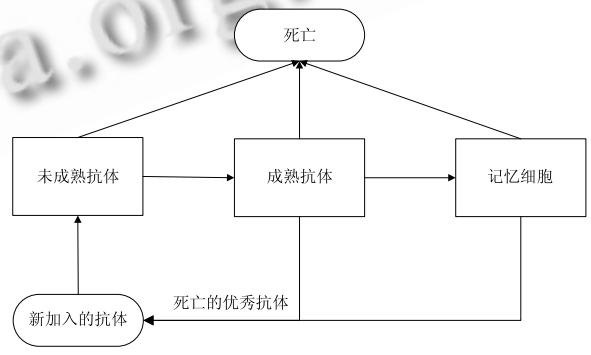

图 1 抗体的生命周期

4 安全工作位置选择的免疫算法

在迁移实例旅行过程中，如果迁移实例在某个工作位置上工作完成，或者该工作位置的服务无法满足迁移实例的需求时，迁移实例就会提出迁移请求。在本文中，迁移实例的迁移目标是有免疫模块规划得到。

免疫算法的设计将在抗原和抗体的编码、亲和度的计算、对抗体信息的维护等 3 个方面进行。算法设计的基本步骤如下：

步骤 1：迁移实例提出迁移请求，并提交任务-服务关系(由业务过程定义)给免疫模块。

步骤 2：免疫模块分析步骤 1 所提交的任务-服务关系，确定任务的资源服务需求，以及所涉及的服务操作，由这些操作确定可能会发生的攻击类型，并确定这些攻击的危险级别。

步骤 3：根据步骤 2 的分析对服务需求和安全信息进行编码，形成抗原，抗原包含任务需求信息以及安全信息。

步骤 4：由步骤 3 所生成的抗原，并根据(1)式计算抗体库中与抗体的匹配程度，即亲和度，亲和度是评价一个抗体优劣的标准。如果存在亲和度符合要求的抗体，则执行步骤 5；否则，需要从服务引擎中随机选取满足服务需求的工作位置，生成新抗体加入到抗体库中，并重复步骤 4。

步骤 5：从步骤 4 所提交的符合要求的抗体中，选取最佳部分抗体，并探测这些抗体所对应的工作位置当前是否可用。

步骤 6：选取可用的最佳工作位置，并让迁移实例进行迁移。

步骤 7：若迁移成功，则反馈免疫模块，由免疫模块更新抗体信息，包括抗体对应的工作位置的服务信息以及抗体生命力等。若迁移未成功，则需重新执行步骤 4-步骤 6，直到迁移成功。

在迁移实例迁移成功以后，免疫模块需要对抗体库进行维护，维护的准则遵循抗体的演化，包括抗体的类型变迁、生命力更新、工作位置信息更新和抗体的死亡等等，这样就使得抗体的信息能够准确的反映当前状态下工作位置的信息。具体的算法流程图如图 2 所示。

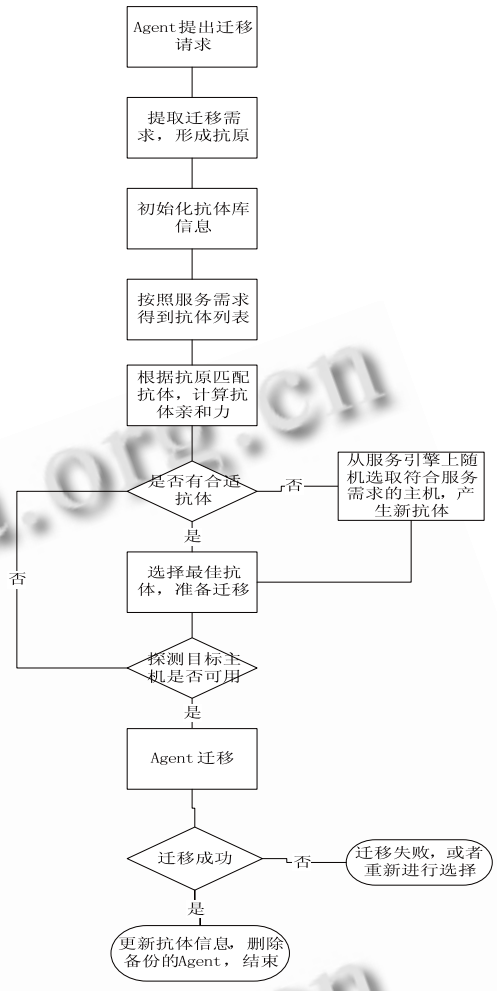

图 2　基于免疫算法的安全工作位置选择算法流程图

5　仿真实验与分析

模拟一个迁移实例的任务序列为 T=(S1,S2,S3,S4,S5,S1,S2,S3,S4,S5)，其中 Si 为迁移实例所需的服务(为简化实验，只考虑服务的数量)，一共有 20 个工作位置，分为 5 组，每组提供一个不同的服务，具体为：S1(W1,W2,W3,W4)，S2(W5,W6,W7,W8)，S3(W9,W10,W11,W12),S4(W13,W14,W15,W16),S5(W17,W18,W19,W20)。为了增大迁移实例的被攻击概率，实验中设置每组中有两个恶意主机，迁移实例在恶意主机上工作时会受到攻击，模拟的攻击类型共有 5 种：拒绝服务、窃取信息、篡改信息、欺骗服务、其他攻击。

仿真实验采用了两种方法，一种是本文提出的基于免疫算法的工作位置选择算法，另一种是一般的选择算法(只考虑工作位置的服务)。依据文中抗原和抗体的编码方式，仿真实验中某个抗原编码方式为(a1，(S1,100)，(001,0.05))，表示标识为 a1 的迁移实例当前任务需要服务 S1，服务数量约束为 100，该服务可能遭受攻击的代码为 001，危险级别为高(0.05)；与之相对应的一个符合要求的抗体的编码方式为(w3，(S1,300)，[(001,0.25), (010,0.01)]，mature，90%，15)，表示工作位置 w3 提供 S1 服务，该服务剩余数量为 300，该抗体对应的工作位置有两种违规类型：001 和

010，抗体类型为成熟细胞，服务成功率为 90%，生命力为 15。

仿真实验模拟迁移实例携带任务说明书一共出行 100 次，迁移实例在每次旅行中的被攻击概率如图 3 所示：

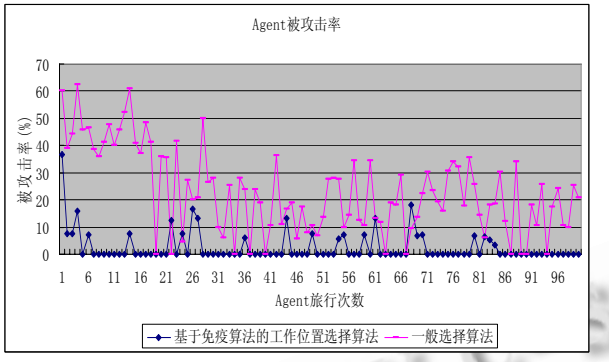

图 3 迁移实例被攻击率变化曲线

上述实验的模拟虽然有夸大的部分(实际环境中恶意主机的比例不会太高)，但从图 3 中可以看出，本文提出的基于免疫算法的工作位置选择算法的被攻击率远远小于一般的选择算法，从而证明了本算法是可行的，对迁移实例安全起到了一定的保护作用。但该方法在初始的工作环境下无法避免恶意主机的攻击，因为初始时，迁移实例所携带的抗体库信息并不完备，需要工作环境相对稳定，即需要迁移实例旅行一段时间，抗体库中信息相对完备的情况下，才能大大的减少迁移实例的被攻击概率，如何有效解决该问题是我们进一步要做的工作。

6 总结

本文提出了一种基于免疫算法的工作位置选择算法，不仅考虑迁移实例自身的任务需求，而且将迁移实例所携带任务的风险等级以及工作位置的安全性考虑在内，从而动态的规划得到相对安全迁移位置。仿真实验表明，该算法有效的避开了恶意主机，使迁移实例的被攻击率大大降低，保护了迁移实例的安全。

参考文献

1 曾广周,党研.基于移动计算范型的迁移工作流研究.计算机学报,2003,26(10):1343−1349.

2 张云勇,刘锦德.移动 Agent 技术.北京:清华大学出版社,2003.

3 曹大军,徐贤良,等.移动 Agent 框架的设计及其关键技术.计算机工程与应用,2002,(1):165−166.

4 何炎祥,陈莘萌.Agent 和多 Agent 系统的设计与应用.武汉:武汉大学出版,2001.

5 Brewington B, Gray R, Moizumi K. Mobile agents in distributed information retrieval. In: Klusch M, ed. Intelligence Information Agent. Berlin: Springer-Verlag, 1999. 355−395.

6 刘大有,杨博,杨琨,王生生.基于旅行图的移动 Agent 迁移策略.计算机研究与发展,2003,(6):838−845.

7 党辰,王嘉祯,等.一种动态环境下的移动 Agent 智能迁移算法.计算机工程,2009,35(9):19−21.

8 张正球,蔡声镇,余敏.一种改进的基于迁移计划图的移动 Agent 迁移策略.计算机应用研究,2007,(1):40−45.

9 李鲁艳,曾广周.基于任务片的旅行图生成算法研究.计算机工程与应用,2008,44(32):41−44.

10 Dasgupta D. Artificial Immune Systems and Their Applications. Berlin Heidelberg: Springer-Verlang, 1999.

11 Morik T, Fukudat. Application of an immune algorithm to multi-optimization problems. Electrical Engineering in Japan, 1998,122(2):30−37.

12 Huang S. An immune-based optimization method to capacitor placement in a radial distribution system. IEEE Trans. on Power Delivery, 2000,15(2):744−749.

13 宁黎华,古天龙.基于免疫算法的装备序列规划问题求解.计算机集成制造系统,2007,13(1):82−87.