基于免疫原理的多代理网格入侵检测系统体系结构研究*

倪建成[1,2], 李志蜀[2], 陈 黎[2], 叶 军[2]
(1. 曲阜师范大学 计算机科学学院, 山东 日照 276826; 2. 四川大学 计算机学院, 成都 610065)

摘 要: 为设计具有高检测率和低误报率的网格实时入侵检测系统, 借鉴人类免疫系统与网格安全系统的相似性特征, 在建立体系结构设计原则的基础上, 确立了群体 agent 实现网格入侵检测系统的交互决策机制; 在选定免疫模型与免疫算法的基础上, 给出了人类免疫系统与网格入侵检测系统组件间的具体映射关系, 从而实现了概念的逻辑移植; 在建立 agent 网格体系结构基础上, 设计了网格入侵检测系统的具体模块框架, 详细阐释了与模块对应的具体 agent 结构与功能, 从而完成了逻辑实体的物理移植。体系结构的建立为设计功能完善的网格入侵检测系统提供了一种可行的逻辑方案。

关键词: 代理; 体系结构; 网格; 免疫; 入侵检测系统
中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2009)02-0705-04

Research on architecture of grid intrusion detection system based on immunity and multi-agents

NI Jian-cheng[1,2], LI Zhi-shu[2], CHEN Li[2], YE Jun[2]
(1. *School of Computer Science, Qufu Normal University, Rizhao Shandong* 276826, *China*; 2. *School of Computer Science, Sichuan University, Chengdu* 610065, *China*)

Abstract: In order to design real-time intrusion detection system for grid environment having high true positive rate and low false positive rate, interactive decision-making mechanism inspired from similar characteristics containing between human immune system and grid security system, was confirmed for grid intrusion detection system implemented with colony agents, which were guided by several established design principles of system architecture. Thereafter, on the basis of the selected immune model and immune algorithms, logical transplant of concepts was solved in describing the concrete mapping relationship between components of human immune system and grid intrusion detection system. Physical transplant of logical entities, and then, was fulfilled by introducing agent grid architecture, designing modular framework of system, and explaining the structure and function of corresponding agents in correlative module. Therefore, a feasible logical scheme was prorvided to construct practical grid intrusion detection system by establishing the detailed system architecture.

Key words: agent; architecture; grid; immunity; intrusion detection system

网格[1]是一个由多用户、多信任组织组成的复杂系统, 具有动态、自组织、多样性和多层等特征, 而这些特征与人类免疫系统(human immune system, HIS)[2]具有惊人的相似性。但是, 它们之间也存在一些不容忽视的差异: a) 在物理层次上, 人类免疫响应是基于组织环境的局部响应, 细胞之间的信息交互依赖于细胞因子等生物化学信号; b) 网格信任组织则是一个更加松散的资源与用户的动态集合, 组织内站点之间的信息交互则是通过网络进行传播, 因而必须具备更高的可靠性和安全性保障。在逻辑层次上, 概念的不一致需要首先解决 HIS 与集成先天性免疫和适应性免疫 agents 的网格入侵检测系统(grid intrusion detection system based on integrated innate and adaptive immune agents, GIDSIIA)组件间的对应关系, 才能将 HIS 安全移植到网格环境中。

体系结构的建立对网格入侵检测系统的具体设计与实现具有重要的指导作用。它一方面在构造方法学上为详细设计 GIDSIIA 组件的演化与交互提供了逻辑框架; 另一方面, 在系统移植上为人类免疫系统到网格环境的移植奠定了物理基础。

1 设计原则

类似于人类免疫系统, 网格入侵检测系统(GIDS)是保护网格软件、硬件以及信息资源的最后一道防线, 其地位在网格安全体系结构中至关重要。为了满足网格的灵活性、分布性、可缩放性和安全高度可控的需求, GIDS 组件必须满足一定的设计原则, 而这些原则可以借鉴危险理论模型中人类免疫系统的基本工作原理。文献[3]仅给出了适用于免疫 agents, 而非人工免疫系统整体的九条基本原则。因此, GIDSIIA 的设计原

收稿日期: 2008-05-07; 修回日期: 2008-07-26 基金项目: 国家自然科学基金资助项目(60072014); 四川省科技公关资助项目(05GG021-003-2); 山东省自然科学基金资助项目(Q99G03)

作者简介: 倪建成(1971-), 男, 山东曲阜人, 副教授, 博士, 主要研究方向为网格计算、人工免疫(nijch@163.com); 李志蜀(1946-), 男, 重庆人, 教授, 博导, 主要研究方向为计算机网络、智能控制与多媒体技术等; 陈黎(1977-), 女, 重庆人, 博士, 主要研究方向为计算机网络与信息系统; 叶军(1978-), 男, 浙江绍兴人, 博士研究生, 主要研究方向为网络与信息系统、Web 服务、网格计算.

则主要包括以下几点:

a) 多样性。GIDSIIA 包含多类 agents, 即控制/决策/监视/免疫 agents。其中: 控制与决策 agent 各自包括节点和社区 agent 两类; 免疫 agent 分为 DC agent 和 T agent。DC agent 包括未成熟/完全成熟/半成熟 DC agent; T agent 包括未成熟/成熟/记忆 T agent。此外, 监视 agent 根据算法需要可被分为网络/进程/系统资源/用户监视 agents。

b) 交互性。入侵检测任务依赖于 agent 与组织环境以及 agent 间的合作与信息传递。

c) 结构化。每类 agent 都具有不同的数据结构。

d) 行为化。信号体现检测任务的行为特征, 包括外部信号和内部信号两部分。外部信号由入侵或被监视系统产生; 内部信号由 agent 分泌。

e) 生命周期。各类 agent 具有不同的生命周期, 存在产生、发育和死亡等不同过程。

f) 功能化。不同类 agent 具有不同的功能。

g) 分化特征。免疫 agent 依据不同的结构和行为特征具有不同的分化路径。

2 HIS 到 GIDSIIA 的移植

人类免疫系统是由循环系统、免疫器官、免疫组织、免疫细胞和免疫分子组成的具有防御和自我调节机制的多层次系统。其中, 免疫组织是决定免疫响应的控制器。免疫响应包括两种类型: a) 先天性免疫响应。依赖非特异性抗原识别与信息融合功能, 先天性免疫能够激发、调节与控制适应性免疫响应类型和强度。b) 适应性免疫响应。适应性免疫系统则依赖特异抗原识别、免疫记忆、否定选择和克隆等机制识别、移除与钝化入侵病原体。人工免疫系统[4] (artificial immune system, AIS) 是对生物免疫系统的细胞功能模型进行抽象而衍生的算法集合, 因此, 目前存在两种典型的人工免疫模型, 即自体/非自体 (SNS) 与危险理论(DM) 模型。当 AIS 应用于网格入侵检测问题时, 具体免疫原理和组件的选择是必须解决的首要问题。事实上, AIS 作为研究者从不同学科、不同角度对 HIS 抽象出来的相应计算机制, 其本身就依赖于具体的应用领域。因而, 基于设计原则, GIDSIIA 结构设计需要考虑的问题包括 HIS 模型/免疫细胞/算法的选择、HIS 与 GIDSIIA 概念的映射关系等。

2.1 免疫模型的选择

在 SNS 模型中, 自体的完备性是决定免疫耐受的关键因素。当自体集不完备时, 系统会产生误报, 这些误报信息无疑会严重增加系统安全管理员的负担; 即使能在有限时间收集到完备的自体集, 它也会因用户行为的变化、系统安全策略等的改变而产生变异, 从而导致系统的高漏报率。在动态网格环境中, 自体频繁变异使系统难以保障自体集的完备性, 因而, 保障系统误报率和漏报率不因自体集的改变而产生急剧变化便成为选择免疫模型的重要衡量指标。

事实上, 基于 SNS 模型的 HIS 理论认为自体与非自体存在明显的标记特征, 即在任一时刻, 免疫系统均能明确区分它们并对非自体实施免疫响应, 因而, 已完成中央耐受的 T 淋巴细胞在任何情况下都能对与抗原配基匹配的 B 淋巴细胞发出正确的协同刺激信号。但是, 基于危险模型的 HIS 理论则认为组织是免疫响应的最终控制者[5]。组织利用正常状态下的外围免疫耐受或异常状态下的免疫刺激, 通过抑制或激活相应的 T 淋巴细胞实施恰当的免疫响应。危险理论更加强调组织的自我调节能力而非自体/非自体的区别, 所以从根源上解决了自体变异导致的漏报问题以及自体不完备导致的误报问题。

2.2 免疫细胞的选择

在 HIS 中, 树突状细胞(DC) 是一种几近遍布全身所有组织的抗原提呈细胞[6]。在未成熟状态时, 它们利用细胞表面富含的的各类受体(如 TLR) 匹配外部环境中的分子; 成熟后, 它们迁移出组织至二级淋巴器官, 向纯真 T 淋巴细胞提呈 MHC 分子上的缩氨酸。MHC 抗原与纯真 T 淋巴细胞表面受体的匹配是引发免疫响应的第一个步骤, 即产生了第一信号, 但是它不能独立激活纯真 T 细胞。第二信号, 即未成熟 DC 被激活后在其表面生成的反映组织环境危险程度的协同刺激分子, 与 T 细胞的交互才能激活纯真 T 细胞并导致其分化。最终, Th2 辅助细胞为 B 淋巴细胞提供协同刺激信号并使之产生抵御入侵抗原的免疫响应; Th1 辅助细胞激活自然杀手细胞与细胞毒素 T 细胞并导致变异自体的清除。

为简化 GIDSIIA 设计, 系统仅利用 DC 和 T 淋巴细胞完成免疫功能。其中: DC 执行信息采集、融合与分泌细胞因子等功能; T 细胞则融合了 HIS 中 T 与 B 细胞的免疫检测功能。

2.3 免疫算法的选择

在 HIS 中, DC 主要完成两方面的功能: 针对抗原, DC 具有抗原采集、处理和提呈功能; 针对信号, DC 具备采集、融合和分泌功能。因而, 在 GIDSIIA 中利用 DCA 算法[7] 融合多层次危险信号以形成与时间相关的抗原相关上下文, 利用适应性免疫系统的否定选择算法避免由基因库形成的免疫细胞产生自我免疫, 利用危险基因生成算法[8] 构造基因库以降低免疫细胞生成时的时间与空间复杂性, 从而提高系统的可缩放能力。

2.4 HIS 与 GIDSIIA 间的组件映射

从构成元素上而言, 种群、生物个体、子系统、器官、组织、细胞和分子组成了具有集合包含关系的人类。与之类似, 网格由不同类型的网格组成, 每种网格由不同的虚拟组织构成, 而不同虚拟组织则包括不同网段的站点资源, 每个节点由位于不同硬、软件系统之上的服务组成。它们之间的映射关系如表 1 所示。

表 1 HIS 与 GIDSIIA 之间的组件对应关系

HIS	GIDSIIA
人类	网格
种群	计算网格
社区	虚拟组织
个体	网段
器官	站点
组织	主机
细胞	进程
分子	网络数据包、系统调用、资源利用率等
抗体	T 免疫细胞
抗原	网络数据包、系统调用、会话连接等

3　整体体系结构

从软件工程的角度而言, 系统的整体体系结构设计是详细设计的基石。对于复杂的网格系统, 单一的安全技术、措施和设备已经不能完全保证网格资源的安全, 因而, 在有效利用现有安全设备、协议、服务等的基础上, 开发与部署适应网格环境的入侵检测系统是必要的防御措施。

3.1　Agent 网格体系结构

Agent 网格[9] 通过集成各种分布式、异质软件组件为上层应用程序和终端用户提供各类网格服务, 提供了一种快速构建应用程序及动态配置网格的方法。图 1 描述了抽象的 agent 网格层次架构。

计算网格层负责集成分布式计算资源以解决超级计算、协同计算和需求计算任务; 数据和信息网格层负责集成网格节点的各类数据类型的信息, 并以统一的形式向上层网格服务提供查询功能; 作为数据与信息网格层的扩展, 网格服务层利用相关软件组件提供资源和数据访问的统一接口。其中: 网格安全基础设施中的认证、授权、代理、加密等服务用于保障 agent 间的通信安全; agent 网格层是用于完成特定任务的各种智能软件模块; 应用程序层利用 agent 网格层的各类(组) agent 提供的服务功能完成用户的需求。

Agent 网格层是整个架构的核心模块, 包括各类 agent 和 agent 系统。Agent 系统是 agent 的运行平台, 负责 agent 功能的实施。单一 agent 用于实现管理、协作、审计、检测等功能; agent群组为实现某一功能的单一 agent 集合。

3.2　GIDS IIA 模型模块体系结构

建立于 agent 网格体系结构之上的 GIDSIIA 模型, 逻辑上包含多个动态形成的信任社区, 每个信任社区包含多个成员节点, 如图 2 所示。各成员节点负责实现入侵检测、分析、响应和提供用户接口, 包括成员节点控制模块(NCM) 、入侵检测模块(IDM) 、入侵决策模块(DMM) 、资源监视模块(RM) 和入侵响应模块(PM) 。其中, IDM 又分为先天性免疫模块(IIM) 和适应性免疫模块(AIM) 。每个信任社区负责监视、控制成员节点的安全状态、接收和发送疫苗等, 包括社区控制模块(TCM) 和疫苗形成与发送模块(VFDM) 。

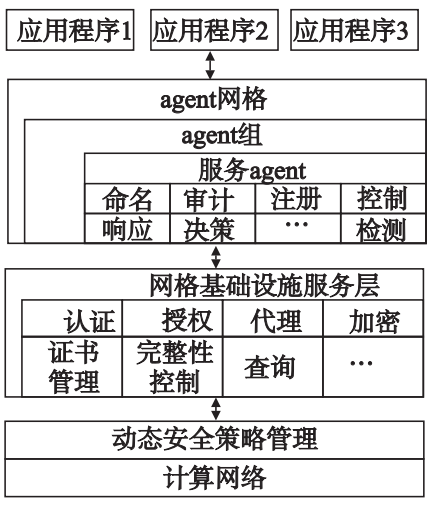

图1　Agent网格层次架构

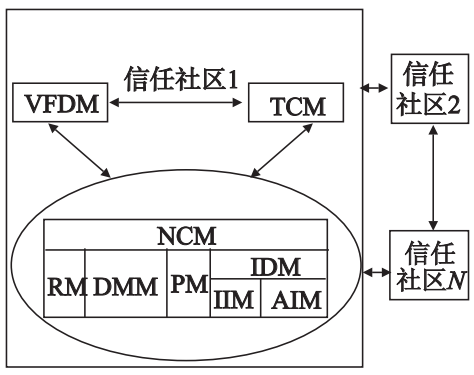

图2　GIDSIIA模型整体体系结构

IDM 模块是整个系统的核心模块, 依据 RM 模块利用事件激发机制收集的成员节点的有关系统状态信息、网络传输信息和系统调用信息等实施具体的入侵检测任务。IIM 主要利用 DC 的信号融合功能快速判断系统的安全状态, 并依据生成的细胞因子激发与控制适应性免疫系统的响应。AIM 主要利用免疫细胞的特异性抗原识别功能, 检测具体的攻击行为。依据 IDM 的具体入侵检测结果, DMM 自动生成相应的防御策略并传递给 PM 具体执行。此外, 对严重的攻击还将形成传递给 VFDM 模块的决策支持信息。

在信任社区层, VFDM 模块集成成员节点的本地决策支持信息形成疫苗, 并依据安全策略在成员节点和信任社区间共享。TCM 与成员节点的 NCM 利用 agent 的心跳信息完成各类 agent 的生成、控制、状态的监视。TCM 与 NCM 存活状态信息的交换极大增强了系统的鲁棒性。

4　Agent 体系结构

在单个成员节点, GIDSIIA 模拟 HIS 的群体 agent 协同工作机制检测、分析和响应入侵; 在信任社区内(间) , GIDSIIA 模拟 HIS 的疫苗分发及接种机制实施入侵信息的传播, 以有效提高网络间的协同检测和二次响应能力。

如图 3 所示, GIDSIIA 包含了多种类型的 agent, 它们都具有不同的功能、结构和行为。在信任社区和成员节点层, 各类 agent 通过信息的交互协同完成入侵检测、信息聚集与关联和疫苗接种等任务。表 2 给出了 GIDSIIA 中的组件模块与相应 agent 间的对应关系。

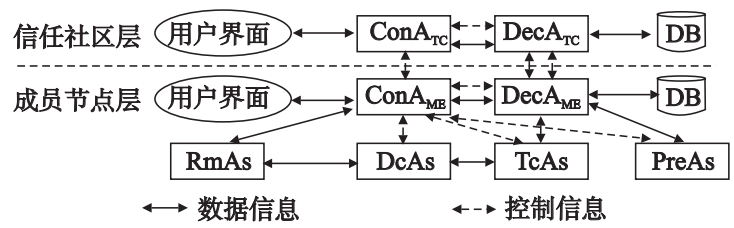

图3　GIDSIIA 信任社区内agent 体系结构

表 2　GIDSIIA 模块组件与 agent 的对应关系

组件名称	agent 名称	组件名称	agent 名称
TCM	ConA TC	DMM	DecA ME
VFDM	DecA TC	RM	RmA
NCM	ConA ME	PM	PreA
IDM	TcA, DcA		

4.1　资源监视 agent

RmA 是数据收集器, 负责系统、网络、进程和用户四个层次相关资源的监视。根据入侵检测算法的需求, 系统层次资源包括 CPU 的利用率、主存资源的使用量和系统文件的属性等; 网络层次资源包括网络数据包的统计信息, 如源、目的 IP 地址和端口号、协议类型、标志位状态及其他流量统计信息等; 进程层资源包括进程名、PID、系统调用序列等; 用户层资源包括用户名、用户 ID、会话连接数等。此外, 收集的信息被加盖时间戳后将形成组织的外部信号和抗原源数据。

4.2　树突状细胞 agent

在 HIS 中, DC 作为先天性免疫响应系统的重要组件, 拥有复杂的组成结构和信息处理功能。DC 利用模式识别受体(PRR) 识别由进化压力形成的与危险或有害病原体有关的抗原相关分子模式(PAMP) , 感知异常出现时的多层次危险信号; 进而绑定抗原与系统环境信息形成与时间有关的抗原上下文信息, 即 DC 将要提呈的抗原集、标志 DC 成熟度及可靠程度的协同刺激分子和调节免疫响应程度的完全成熟/半成熟细胞因子。

DcA 模拟树突状细胞的功能, 实现危险信号的感知、抗原的捕获、抗原相关危险信号的融合以及 DC 的分化, 包括未成

熟/完全成熟/半成熟三类 DcA。RmA 传送的病原体基因片段和多层次内外部信号被虚拟组织接收和预处理后,分别存储并构成病原体基因片段集与内、外信号集。未成熟 DcA 首先从上述三类信息集中分别提取抗原及相关信号形成不同的数据子集;既而进行抗原特征提取及合成、抗原与相关内/外部信号的上下文信息融合;最后产生类似 MHC 分子的抗原集与其所处系统环境的上下文信息并分化为完全成熟 DcA 或半成熟 DcA。代表不同攻击强度和可靠性的完全成熟/半成熟 DC agent携带与时间有关的上下文相关抗原迁移至淋巴节点,进而激发和控制 TcAs 进行具体攻击的识别与判断。

4.3 T 细胞 agent

在 HIS 中,适应性免疫系统主要由 T 淋巴细胞和 B 淋巴细胞进行适应性免疫响应。B 细胞的主要功能包括识别抗原和分泌抗体。当生物体遭受细菌、病毒等外部入侵时,B 细胞表面的 BCR 识别其 MHC-II 分子,亲和力成熟并被激活后,该 B 细胞将分化为能分泌抗体的浆细胞,然后浆细胞利用分泌的抗体绑定入侵病原体并进一步使其他细胞杀死、移除被绑定物质。T 细胞包括 $CD4^+$ 和 $CD8^+$ 两个子类,主要功能为调节其他细胞的行为和直接杀死被感染细胞。$CD4^+$ T 通常作为辅助细胞,依据分泌的细胞因子,可进一步分为 Th1、Th2 和 Tr(抑制性 T 细胞)。Th 细胞识别被 MHC-II 分子提呈的抗原肽,并调节细胞和体液的免疫响应[10]。Th1 细胞主要用来刺激 CTL 细胞、NK 细胞和巨噬细胞以消除细胞内病原体;Th2 细胞负责刺激 B 细胞分泌抗体以对付外部病原体的攻击;Tr 具有潜在的免疫抑制作用。$CD8^+$ T 细胞作为细胞毒素,通常用来识别 MHC-I 分子提呈的抗原肽。当 $CD8^+$ T 细胞被激活后便成熟为细胞毒素 T 细胞(CTL),用来杀死被细胞内细菌或病毒感染的变异自体细胞。

TcA 模拟 T 淋巴细胞和 B 淋巴细胞的功能,实现抗原的匹配、与 DC 的交互、免疫响应和细胞的分化,包括未成熟/成熟和记忆 TcA。完全成熟 DcA 与记忆 TcA 的肯定交互引发快速的二次响应,而与成熟 TcA 的肯定交互将在抗原匹配数超过激活阈值时使其分化为记忆 TcA 并产生首次免疫响应;半成熟 DcA 与记忆/成熟 TcA 的肯定交互则钝化并删除匹配的特异性抗原,避免误报的产生。当产生免疫响应时,相关的进程、系统等信息发送至决策 agent 作进一步的信息聚集与关联。

4.4 决策 agent

决策 agent 分为成员节点决策 agent($DecA_{ME}$)和信任社区决策 agent($DecA_{TC}$)。成员节点决策 agent 负责接收本地节点 TcA 的攻击信息,通过聚集与关联形成更高层次的攻击模式(发送模式信息至 $DecA_{TC}$),并依据攻击的危险程度制定具体的响应策略(发送响应措施至 PreA)。此外,成员节点决策 agent还接收社区决策agent发送的疫苗。类似成员节点决策 agent,信任社区决策 agent 负责聚集与关联社区内各节点 $DecA_{ME}$形成的攻击信息,产生相应的疫苗后,依据既定安全策略将其在社区内和社区间进行分发,进而使整个网格环境具有快速响应入侵的能力。

4.5 响应 agent

在 HIS 中,免疫系统拥有多种复杂的响应机制进行病原体的清除和调节免疫系统的平衡。例如,CTL 细胞、NK 细胞和巨噬细胞等可以消除细胞内病原体;抗体可以对付外部病原体的攻击;Tr 具有潜在的免疫抑制作用等。

入侵响应系统根据不同的攻击类型可以采用报警、自动响应和人工响应等方式响应入侵。在 GIDSIIA 中,响应 agent(PreA)从 $DecA_{ME}$接收防御信息后,根据既定响应策略可以采用自动延迟系统调用的响应方式。

4.6 控制 agent

控制 agent(ConA)是系统的重要控制单元,与系统内的其他 agent 通过周期性的心跳信息感知彼此的存活状态,提供系统管理员与 GIDSIIA 交互的接口,控制其他 agent 的启动和停止。此外,成员节点控制 agent($ConA_{ME}$)与信任社区控制 agent($ConA_{TC}$)之间以及 $ConA_{TC}$与 $DecA_{TC}$之间组成多对多的映射关系。

5 结束语

网格环境的安全保障是一个富有挑战意义的议题。本文借助 HIS 与网格安全系统组件功能与行为特征的相似性,建立了网格入侵检测系统基本设计原则,确立了 GIDSIIA 的设计指导思想,即采用多样化的具有独立功能、结构和行为的移动 agent 作为基本组件,利用群体 agent 交互机制实现入侵检测和响应功能。然后在选取免疫模型与免疫算法的基础上,给出了 HIS 与 GIDSIIA 组件间的具体映射关系,进而在建立基于 agent 网格的整体体系结构基础上,设计了 GIDSIIA 的具体模块交互框架,并详细说明了与模块对应的具体 agent 结构与功能,从而实现了 GIDSIIA 模型从 HIS 的移植。

Agent 体系结构是描述 agent 演化过程的基础,未来的工作将集中于 agent 功能与行为的实现研究。

参考文献:

[1] OSEPH J, FELLENSTEIN C. Grid computing [M]. 战晓苏,张少华,译. 北京:清华大学出版社,2005.

[2] JANEWAY C A, TRAVERS P, WALPORT M. Immunobiology[M]. 5th ed. New York; London: Garland Science, 2001.

[3] TWYCROSS J. Integrated innate and adaptive artificial immune systems applied to process anomaly detection[D]. Nottingham: Nottingham University, 2007.

[4] FORREST S, HOFMEYR S A, SOMAYAJI A. Computer immunology[J]. Communications of the ACM, 1997, 40(10): 88-96.

[5] MATZINGER P. The danger model: a renewed sense of self[J]. Science, 2002, 296(5566): 301-305.

[6] ESCHE C, SHURIN M R, LOTZE M T. The use of dendritic cells for cancer vaccination[J]. Current Opinion in Molecular Therapeutics, 1999, 1(1): 72-81.

[7] GREENSMITH J. The dendritic cell algorithm[D]. Nottingham: Nottingham University, 2007.

[8] LQBAL A. Danger theory metaphor in artificial immune system for system call data[D]. Jhelum: Mohallah Islamia High School, 2006.

[9] CHI Jia-yu, SUN Ling, CHEN Xue-Guang, *et al*. Architecture and design of distributed decision support system based on agent grid [C]//Proc of the 4th International Conference on Machine Learning and Cybernetics, 2005: 321-327.

[10] DELVES PJ, ROITT IM. The immune system: second of two parts [J]. The New England Journal of Medicine, 2000, 343(11): 108-117.