网络空间拟态防御研究

邬江兴

国家数字交换系统工程技术研究中心 郑州 中国 450001

摘要 本文扼要分析了网络安全不平衡现状及本源问题, 重点阐述了动态异构冗余架构以及如何利用不可信软硬构件组成高可靠、高安全等级信息系统的原理与方法, 概略的给出了拟态防御的基本概念。最后, 介绍了拟态防御原理验证系统的测试评估情况和初步结论。

关键词 网络安全; 不平衡态势; 未知漏洞后门; 拟态防御; 非相似余度; 动态异构冗余; 验证系统测试评估
中图法分类号 TP309.1 DOI 号 10.19363/j.cnki.cn10-1380/tn.2016.04.001

Research on Cyber Mimic Defense

WU Jiangxing

National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450001, China

Abstract The unbalance in cyber security and its root is analyzed in this paper, of which a dynamically redundant heterogeneity architecture and the basic principle to compose an information system of high reliability and high security level with unreliable hardware is mainly discussed. Then, basic concepts about mimic defense is briefly stated. Finally, a test evaluation and preliminary conclusion about the system of mimic defense verification is presented.

Key words Cyber security; unbalance; unknown bug and backdoor; mimic defense; dissimilar redundancy; dynamically redundant heterogeneity; test and evaluation

1 背景

正当人们尽情享受信息时代生活和工作乐趣的同时, 网络安全问题像幽灵一般成为挥之不去的梦魇。根本原因之一是, 网络空间存在泛在化的基于未知漏洞和后门等的不确定性威胁。这使得少数人或团体组织利用少量的资源就能肆意践踏个人乃至公众的隐私权, 威胁信息基础设施和公共服务系统安全, 危及网络空间秩序和社会稳定。当前, 网络空间的基本安全态势是“易攻难守”。

1.1 网络安全不平衡态势的本源问题

首先是未知安全漏洞问题。通常是指, 在硬件、软件或协议等的具体实现或系统安全策略上存在的缺陷, 从而使攻击者有可能在未经授权下访问或破坏系统。由于人们认知方法与手段的局限性, 实践中, 设计缺陷或安全漏洞往往难以避免。

其次是软硬件后门问题。通常是指留在软硬件系统中的恶意代码, 旨在为特殊使用者通过特殊方式绕过安全控制环节而获得系统访问权提供方法与途径。“棱镜门”事件披露了大量的关于后门的黑幕消息, 给全球信息技术产品市场带来了极为严重的负面影响。即使在网络安全技术占主导地位, 高端信息产品市场占绝对优势的美国, 2017 年国防授权法案中也提出了“如何保证来自全球化市场、商用等级、非可信源构件的可信性问题”。供应链全球化时代, 形成了“你中有我, 我中有你, 相互依存又互为掣肘”的生态关系。因而, 在可以预见的将来, 后门问题也是无法杜绝的。

排在第三位的是未知漏洞后门或侧信道的发现问题。迄今为止, 人类尚未形成穷尽复杂信息系统漏洞和彻查后门的理论与方法。就目前的科技能力来看, 网络空间能够查出的漏洞和后门就如同浩瀚宇宙中的一点点尘埃, 绝大多数是未被认知的。更为糟糕的是, 侧信道带来的安全隐患常常是泛在化的且很难用物理或逻辑的方法彻底消除。在“设计缺陷与不可信的开放式供应链”条件下, 无论从技术或经济上都不可能保证网络空间构成环境内“无毒无菌”的理想安全性要求。

从“已知的未知”风险防范视角来看, 基于未知

邀稿日期: 2016-05-06; 收稿日期: 2016-09-26
本文工作受到国家自然科学基金创新研究群体项目(No.61521003)和国家重点研发计划项目(Nos. 2016YFB0800100, 2016YFB0800101)支持。

漏洞或未知后门实施的未知攻击属于“未知的未知”安全威胁, 即不确定性威胁, 也是网络安全领域最令人惊恐不安、备受煎熬的心理折磨。由于无法适时的了解攻击者采用什么手段、通过何种途径、利用什么缺陷或“内应”, 是否已进入目标对象, 已经干了什么, 正在干什么, 未来打算干什么等等一系列不确定性问题, 因而也就无从谈起怎样才能实施有效的针对性防御。参照经济学理论[1]的相关说法, 已知的未知属于风险, 风险可以用概率来表述, 而未知的未知属于不确定性, 不确定性则是不知道概率的情形。因此, 不确定性威胁是造成网络安全不平衡态势难以逆转的核心因素之一。

1.2 现有防御体系的脆弱性和安全黑洞

现有的防御体系是基于威胁特征感知的精确防御, 需要获得攻击来源、攻击特征、攻击途径、攻击行为和攻击机制等先验知识作为实施有效防御的基础。因此, 它必须建立在“已知风险”或是“已知的未知风险”前提条件上。不幸的是, 由于软硬构件中存在未知的漏洞或后门, 信息系统和防御体系只能构建在可信性无法保证的运行环境上。

从严格意义上说, 现有信息系统或防御体系对泛在的不确定性威胁基本上是不设防的, 除了加密认证外几乎没有其他实时高效的应对措施。事实上, 确保加密认证装置实现上的可信性本身就极富挑战性。如果作为嵌入式装置在非可信环境下使用, 很可能存在被“短路”的风险或受到基于“侧信道”原理的各种旁路攻击[2, 3]。

由于无法保证信息系统或网络空间生态环境的“无漏洞无后门”或者“无毒无菌”, 现有的安全防护体系只能期待“后天获得性免疫”。通过不断地亡羊, 不停地补牢, 不断地挖掘漏洞和发现后门, 不停地打补丁, 杀毒灭马, 封门堵漏等被动的跟随博弈方式来自我完善。但面对不确定性威胁时, 被动防御就如同数学上求解缺维方程组, 理论上无确定解。所以说, 目前网络安全“易攻难守”的不对称态势也是被动防御理论体系和技术的基因缺陷所致。

更为严峻的是, 网络空间信息系统架构和防御体系本质上说都是“静态的、相似的和确定的”, 体系架构透明、处理空间单一, 缺乏多样性。软件的遗产继承将导致安全链难以闭合, 系统缺陷和脆弱性持续暴露且易于攻击, 使之成为了网络空间最大的安全黑洞。

2 相关概念

2.1 拟态现象与拟态防御

一种生物在色彩、纹理和形状等特征上模拟另一种生物或环境, 从而使一方或双方受益的生态适应现象, 在生物学中称为拟态现象[4]。按防御行为分类可将其列入基于内生机理的主动防御范畴, 我们将其称为“拟态伪装”(Mimic Guise, MG)。

拟态现象在生物界其实很普遍, 诸如竹节虫、枯叶蝶、模拟兰花、树叶虫等林林总总, 光怪陆离。尤其是 1998 年在印尼苏拉威西岛水域发现的条纹章鱼(又名拟态章鱼), 算是生物界的拟态伪装大师。据研究表明, 它能模拟 15 种以上海洋生物, 可以在珊瑚礁环境和沙质海底完全隐身。在本征功能不变条件下, 能以不确定的色彩、纹理、形状和行为变化给掠食者或捕食目标造成认知困境[4]。

拟态伪装不能直接作为网络空间主动防御的概念基础。根本原因是, 网络空间中大多数信息系统的服务功能和性能是不能被隐匿的, 如 Web 服务、路由交换、文件存储、云计算和数据中心等等。相反, 还要尽可能让外界清晰明了其使用方式和功能细节以及考虑使用习惯的延续问题。所以, 网络空间的任何主动防御举措都不能影响到目标对象给定服务功能和性能的正常提供。除此之外, 其他行为, 比如目标对象的系统架构、运行机制、核心算法、异常表现以及可能存在的未知漏洞或后门等等, 都可以通过类似拟态伪装的方式进行主动隐匿。那么, 除隐匿目标对象服务功能外的拟态伪装就定义为“拟态防御”(Mimic Defense, MD)。

2.2 拟态防御与非相似余度构造

根据“给定功能, 往往存在多种实现结构”的公理, 可靠性领域的异构冗余体制可能满足拟态防御的最低要求。原因是, 其异构冗余集合中的任意元素, 无论是单独使用还是多个元素的并联或组合使用, 可呈现的功能场景都是等价的。

异构冗余体制的经典范例就是“非相似余度构造”(Dissimilar Redundancy Structure, DRS)[5, 6]。如下图所示:

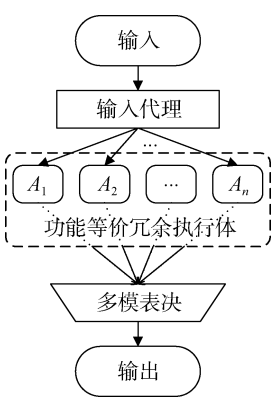

图 1 DRS 结构示意

其理论基础是“独立开发的装置或模块发生共性设计缺陷导致共模故障的情况属于小概率事件”, 工程实现上要保证各功能等价的独立装置中的设计缺陷具有不相互重合的性质。

目前, DRS 系统的相异性设计主要通过严格的管理手段实现。由不同教育背景和工作经历的人员组成的多个工作组, 依托不同的开发环境, 根据不同的技术路线, 遵循一个共同的功能规范, 独立开发多个功能等价的装置。然后, 将这些异构等价装置加入到一个具有多模表决机制的容错架构, 通过对输出矢量的比较或检测, 便可容忍处理可能存在的两种不确定性错误, 即软硬件设计潜在缺陷引发的不确定故障和物理机制导致的不确定失效。这说明, 原本属于不确定性故障的处理难题可以被 DRS 架构转换为择多判决的概率问题。尽管理论上不能保证择多结果总是正确的, 但却可以证明择多错误的发生概率随 DRS 余度数增加呈非线性减小。

工程实践表明, 基于 DRS 构造的系统能够非线性地提高可靠性等级。如果再借助一定的预清洗或恢复机制, 应用系统可以具有非常高的可靠性(例如, 美国波音公司的 B-777 客机飞控系统的失效率就低于 10^{-10}, 通用动力公司的 F-16 战斗机飞控系统的失效率也低于 10^{-8})[7]。值得注意的是, 上述设计原则与方法不仅对抑制目标系统的共性缺陷和错误有效, 而且对管控开发工具和开发环境引入的共性缺陷同样有效。

DRS 的应用成果证明, 异构冗余体制能够有效地检测、定位和隔离“已知的未知”或“未知的未知”故障, 还可以显著减少系统验证和确认的时间与费用[7]。

2.3 DRS 架构抗攻击缺陷

尽管 DRS 架构对于不确定性故障具有很好的容错属性, 但直接用于应对诸如高级持续威胁(APT)等增量性、持续性、协同性和不确定性高级入侵威胁仍存在以下几方面挑战:

1) DRS 的基本要求是所有构件、组件有严格的空间独立性, 且给定功能交集严格限定。除了等价功能交集, 不能存在其他相同的功能交集, 这需要极为复杂、代价高昂的相异性设计来保证;

2) DRS 系统仍然是静态的、确定的, 其运作机制也是相似的, 理论上具有防御行为的可预测性。攻击者一旦掌握目标对象过程信息或所需的必要资源, 就可以从容的造成多数异构体的一致性错误, 以实现多模裁决下的逃逸(例如通过侧信道攻击方式实现跨物理域的协同攻击)。更严重的是, 一旦攻击成功, 其经验具有可继承性, 能够准确的复现;

3) DRS 构造的容错特性是以软硬件故障发生的随机性为前提, 通常可以用概率工具来描述。而基于未知漏洞或后门等发起的攻击对防御方来说属于不确定性威胁, 不具备可计算性质, 也就无法用概率工具进行表达和分析;

4) 虽然 DRS 构造同样基于相对不可靠的组件作为冗余体构件, 但一般认为可靠性隐患不具有“恶意传播”的性质。因此, 并没有强制性的“去协同化”以阻止人为攻击的传播, 冗余体间通常存在协同化攻击可利用的互连端口、链路和处理空间等通道或同步机制。

总之, 一旦攻击者拥有的资源和能力可以覆盖 DRS 架构的静态环境, 理论上就能够实现基于多数冗余体的协同化攻击。

3 动态异构冗余思想

作为一种共同的认知, 在诸如 DRS 等异构冗余体制中导入动态性和随机性可以有效提升其抗攻击性能。首先, 该机制在防御方可控或代价可接受条件下, 能使系统对攻击者呈现出相当程度上的不确定性, 为上节问题(2)提供了强有力的解决手段, 也能将问题(3)中人为攻击行动的必然性强制转化为随机性可表述的事件。其次, 研究经验表明, 使同构冗余部件间做到精确协同极具挑战性, 而动态性和随机性能够进一步提升多方参与具有一致性或协同性要求的行动的不确定度, 多模判决环节又在机理上显著提升了非配合条件下的协同攻击难度。一个自然的结论是, 倘诺异构执行体集合中的元素采取诸如动态化的策略调度, 或是元素自身实施可重构、可重组、可重建、可重定义、虚拟化等广义动态化技术, 那么在保证系统功能、性能及抗攻击与可靠性指标的前提下, 可以简化或弱化 DRS 构架中苛刻的相异性设计要求。此结论对于后全球化时代, 在开放式产业链、供应链环境下, 应用开源软硬件或中间件等不可控构件组成安全可信系统, 降低产品全寿命周期成本, 有着十分重要的工程意义。我们将这样的异构冗余体制称为“动态异构冗余”构造(Dynamic Heterogeneous Redundancy, DHR)。

3.1 DHR 构造

动态异构冗余构造, 理论上要求系统具有视在结构表征的不确定性。包括非周期的从功能等价的异构冗余体池中随机的抽取若干个元素组成当前服务集, 或者重构、重组、重建异构冗余体自身, 或者借助虚拟化技术改变冗余执行体内在的资源配置方

式或视在运行环境, 或者对异构冗余体作预防性或修复性的清洗、初始化等操作, 使攻击者在时空维度上很难有效的再现成功攻击的场景。其架构如下图所示:

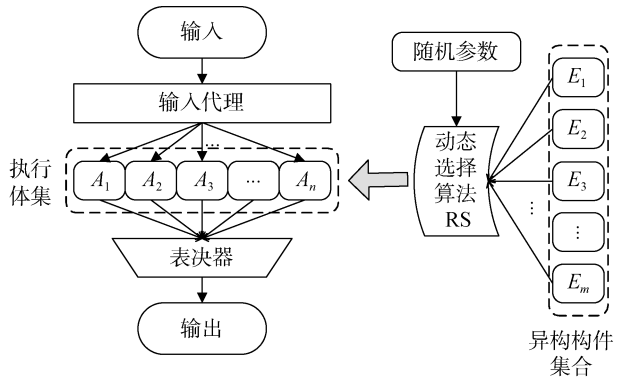

图 2 DHR 结构示意

该构造由输入代理、异构构件集、策略调度算法、执行体集和多模表决器组成。其中, 异构构件集和策略调度组成执行集的多维动态重构支撑环节。由标准化的软硬件模块可以组合出 m 种功能等价的异构构件体集合 E, 按策略调度算法动态的从集合 E 中选出 n 个构件体作为一个执行体集($A_1, A_2, ..., A_n$), 系统输入代理将输入转发给当前服务集中各执行体, 这些执行体的输出矢量提交给表决器进行表决,得到系统输出。我们将输入代理和多模表决器也称为“拟态括号”(Mimic Bracket,MB)。

拟态括号内通常是一个符合 IPO 模型的防护目标之集合(规模或粒度不限), 如下图所示:

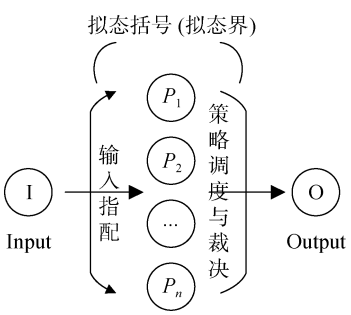

图 3 IPO 模型

P 可以是复杂的软硬件处理系统也可以是部件或子系统, 且存在应用拟态防御架构的技术与经济条件, 则可表达为: $I(P_1, P_2, ..., P_n)O$ 。其中, 连接输入 I 的左括号被赋予输入指配功能, 连接输出 O 的右括号被赋予多模表决和代理输出功能, 括号内的 P_n 是与 P 功能等价的异构执行体。左右括号在逻辑上或空间上一般是独立的, 且功能上联动。拟态括号限定的保护范围称为拟态防御界(Mimic Defense Boundary, MDB), 简称拟态界, 通常情况是一个存在未知漏洞、后门或病毒、木马等软硬件代码的“有毒带菌”异构执行环境。

不难发现, DHR 构造的抗攻击能力在体系上源自 DRS 构造, 在不确定性机制上则受惠于动态性、随机性和多样性的引入, 在攻击难度上得益于“去协同化环境”造就的非配合条件下的协同化攻击困境, 在实现方法上来源于功能等价条件下的多维动态重构机制的应用。

3.2 DHR 构造抗攻击性分析

3.2.1 相关定义

记所有异构构件(简称为构件)组成的集合为 $E = \{E_1, E_2, ..., E_m\}$,构件 $E_i (1 \leqslant i \leqslant m)$ 中所有漏洞组成集合为 V_i ,所有构件中漏洞组成的集合为 $V = \bigcup_{i=1}^{m} V_i$ 。系统各时刻从 m 个构件中选择 n 个构件组成当前系统的执行体集; 所有执行体集构成系统服务集 W, W 中元素的总数目称为系统服务集数 λ , 即 $\lambda = |W| = C_m^n$ 。

定义 1 一次攻击成功

一次攻击成功是指导致系统的机密性、可用性、完整性等遭受破坏的完整攻击过程。本模型中一次攻击指攻击者行为在系统中发生了一次输入输出, 一次攻击成功可能产生多次输入输出, 即进行了多次攻击。

对构件而言, 当攻击者利用构件中的漏洞并成功控制该构件, 则认为构件被攻击成功。对系统而言, 当攻击者利用 n 个执行体中的漏洞发起攻击, 导致某些执行体输出一致但与正常输出不一致, 且异常输出通过了表决器表决, 则认为系统被攻击成功。

定义 2 系统攻击成功率

若攻击者利用 DHR 系统中的漏洞发起 α 次攻击, 导致系统被攻击成功 β , 则记:

$$P_S = \frac{\beta}{\alpha}$$

为一次攻击的系统攻击成功率。

定义 3 系统动态变化时间

给定有限长的系统服务时序 $\Delta t_0 \Delta t_1, \cdots, \Delta t_{k-1}$, 记 $\Delta t_{\min} = \min_{0 \leqslant i \leqslant k-1} \Delta t_i$ 为系统最小变化时间, $\Delta t_{\max} = \max_{0 \leqslant i \leqslant k-1} \Delta t_i$ 为系统最大变化时间, $\overline{\Delta t} = \dfrac{\sum_{i=0}^{k-1} \Delta t_i}{k}$ 为系统平均变化时间。

若系统最小变化时间和最大变化时间相等, 则称系统是周期性变化的。该变化时间称为系统动态变化周期(或系统服务时序周期), 记为 T 。

定义 4 l 阶输出一致率

设 n 个执行体组成一个执行体集, 若在相同输入条件下, 有 l 个输出相同, 且与正常输出不一致, 称该执行体集是 l 阶输出一致的。

考虑到攻击者是针对漏洞进行的, 若对漏洞 v , 在 N 次攻击中有 N' 次其 l 个输出相同且与正常输出不一致, 则称:

$$\varepsilon_l(v)=\frac{N'}{N}$$

是关于漏洞 v 的 l 阶输出一致率。

若忽略系统随机错误, 易知对任意漏洞 v , 关于的 l 阶输出一致率:

$$\varepsilon_l(v)=\frac{\sum\limits_{1\leqslant i_1<i_2<\cdots<i_n\leqslant m}\delta_l(i_1,i_2,\cdots,i_n,v)}{C_m^n}$$

$$\delta_l(i_1,i_2,\cdots,i_n,v)=\begin{cases}1 & \text{if} \quad \sum\limits_{j=1}^{n}\chi_{V_{i_j}}(v)\geqslant l\\ 0 & \text{else}\end{cases}$$

3.2.2 前提假设

由于信息系统的复杂性和网络攻击的多样性, 根据网络攻击的实际情况, 特作如下假设:

(1) 系统的输入与输出是一一对应的。

(2) 针对漏洞的攻击, 若攻击成功必导致执行体输出的改变(与正常输出相比)。

(3) 输入代理和表决器是安全的, 即不考虑输入代理和表决器受攻击的情况。

(4) 攻击者仅基于执行体的漏洞发起攻击, 每次攻击事件是独立的。

(5) 不考虑通信误码或执行体故障等因素引起的系统随机错误。

3.2.3 建模与求解

DHR 系统模型实例可用一个 8 元组 Ω 来表示 $\Omega=\langle m,n,k,T,V,q_v,t_v,\alpha_v\rangle$, 其中:

m : 构件总数, 表示系统有异构构件 $E_1,E_2,\cdots,E_m$;

n : 执行体集大小, 表示每个执行体集含有可变的异构构件 $A_1,A_2,\cdots,A_n$;

T : 动态变化周期, 每个 T 内系统当前执行体集固定;

V : 所有构件中漏洞组成的集合;

q_v :攻击者利用漏洞 v 对含 v 的执行体进行攻击的成功率;

t_v : 攻击者利用漏洞 v 对含 v 的执行体进行攻击所需时间;

α_v : 攻击者利用漏洞 v 攻击的权重;

DHR 系统建模的目的是要揭示系统对抗基于未知漏洞等不确定性威胁的安全机理, 我们定义系统攻击成功率 P_S 来表征系统主动防御的有效性, 并导出其与参数 $m,n,k,T,V,q_v,t_v,\alpha_v$ 的函数关系:

$$P_S=f(m,n,k,T,V,q_v,t_v,\alpha_v)$$

定理 1 给定 $\Omega=\langle m,n,k,T,V,q_v,t_v,\alpha_v\rangle$ 的 DHR 系统, 则一次攻击的系统攻击成功率为:

$$P_S=\sum_{i=1}^{N}\alpha_i\cdot q_{v_i}\cdot\varepsilon_k(v_i)\cdot I(v_i)$$

$$=\frac{\sum\limits_{i=1}^{N}\left[\alpha_i\cdot q_{v_i}\cdot I(v_i)\cdot\sum\limits_{1\leqslant i_1<i_2<\cdots<i_n\leqslant m}\delta_k(i_1,i_2,\cdots,i_n,v_i)\right]}{C_m^n}$$

定理 2 针对特定漏洞 v 的攻击成功率为 $P_v=q_v\cdot\varepsilon_k(v)$, 其中:

$$\varepsilon_k(v)=\frac{\sum\limits_{1\leqslant i_1<i_2<\cdots<i_n\leqslant m}\delta_k(i_1,i_2,\cdots,i_n,v)}{C_m^n}$$

3.2.4 性质分析

(1) $P_v=0$ 当且仅当任意 k 个互异执行体 $A_i(1\leqslant i\leqslant k)$ 无共同漏洞 v 或 $I(v)=0$ 。

这说明当攻击者利用特定漏洞 v 发起攻击时, 欲使系统攻击成功率降低, 必须降低 $P_v=q_v\cdot\varepsilon_k(v)\cdot I(v)$ 值; 由于 q_v 是客观存在的且与攻击者能力密切相关, 因此通过提高系统异构性可以降低 k 阶输出一致率, 也可以通过减小 T 值使得 $I(v)=0$ 。

(2) $P_S=0$ 当且仅当对任意 $1\leqslant i\leqslant N$ 有 $P_{v_i}=0$ 。

这说明欲使系统攻击成功率降低, 必须降低各 P_{v_i} 值。根据分析(1), 系统只能转而控制 T 值或关于各漏洞的 $\varepsilon_k(v_i)$ 值。

这两条性质表明了 DHR 模型对抗攻击者利用漏洞进行攻击的有效性。根据性质(1)系统的动态性使得在任意动态变化周期内攻击时间代价 $t_v\geqslant T$ 的漏洞无法被有效利用, 系统的异构冗余性使得 $\varepsilon_k(v)$ 很小, 攻击者能够有效利用漏洞 v 的时间周期窗口变少且不确定, 从而增大了利用难度。即使在攻击成功情况下, 欲保持攻击效果的持续时间也被压缩到至多 1 个时间周期 T ; 根据性质(2)系统整体上攻击成

功的平均概率减小, 说明其能有效对抗基于漏洞的网络攻击。

3.3 DHR 构造效应

1) 不确定性威胁感知

DHR 构造感知不确定性威胁需要满足三个前提条件: (1)执行体的输入是标准化或可归一化的, 且能被用于并行激励多个执行体; (2)执行体的输出矢量是标准化的或经归一化处理能够满足标准化, 并可支持多模判决; (3)执行体受到的不确定性攻击将在其输出矢量中有所表现。感知精度决定于拟态括号内执行体的颗粒度。

2) 增加攻击链不确定性

DHR 构造内在的多维动态重构机制使得目标对象的视在环境表现出很强的不确定性。理论上除给定功能外, 同样的输入激励很难得到相同的输出响应。例如, 静态情况下攻击者看到的漏洞数量、种类特征与可利用性是确定的, 而在动态情况下视在漏洞的数量、种类特征、出现的频度和可利用性都将是不确定的。这将会严重干扰攻击者在漏洞探测和回连阶段获得信息的真实性, 也会使攻击植入或上传环节失去可靠性, 还将会破坏攻击的成功经验在时空维度上的可复现性。

3) 增加多模裁决逃逸难度

理论上, 多模裁决机制存在“错误逃逸”的可能性, 但逃逸概率通常很低。由于 DHR 构造中导入了策略调度、重构重组和虚拟化等多维动态的不确定性机制, 将使得多模判决环节中单次逃逸或持续逃逸的概率进一步降低。原因是: (1)多模裁决环节对经过拟态界的攻击链频度和执行集的环境差异都十分敏感, 任何导致冗余执行体输出矢量产生差异的功能性操作或内容性设计会被非线性放大; (2)如果从异构冗余池中抽取的执行集是按照某种随机策略进行的, 则试图满足当下服务集内同源组合漏洞的逃逸条件也是相当困难的; (3)假设 t 时刻的攻击实现了成功逃逸, 由于服务集具有动态性和异构性, 在 $t+i$ 的后续攻击中也难以借助逃逸经验复现出与 t 时刻完全相同的输出矢量; (4)如果系统再辅以“发现即清洗”的处理策略, 那么越是复杂精细的攻击方案, 越是操作步骤绵长的协同行动, 攻击成功率就会越低。

4) 动态异构环境影响漏洞利用

原理上, 拟态界内的漏洞只要保证严格的相异性就具有不可利用性, 这是由多模裁决的本身机制决定的。事实上, 即使在拟态界内存在同源或同宗漏洞等情况, 由于执行环境的动态和异构变化, 使得攻击者难以实现针对环境差异的多目标协同攻击。例如, 借助 CPU、OS 或支撑环境漏洞实现的一类攻击将会被多模裁决环节阻断; 利用应用软件漏洞进入系统并设法通过 OS 提权操作的蓄意攻击也将因为环境因素变化而难有作为; 借助冗余体内存、缓存或输入输出等“侧信道”实施的应用层攻击也会被多模裁决机制终止; 根据特定化环境定制的后门、病毒或木马等, 也会随着攻击可达性的改变而失去期望功能。

5) 具有独立的安全增益

DHR 构造的防御机理主要体现在目标对象视在环境与构造的不确定度方面, 目的是非线性的增加攻击者的难度和成本。防御的有效性仅由 DHR 构造的多维动态重构、重组等机制和非配合条件下多模裁决机制的强弱来决定, 不依赖任何攻击行为的先验知识, 也与传统安全防御手段(例如, 入侵检测[8, 9]、预防、容忍[10, 11]或隔离, 杀毒灭马、封门堵漏等)无强关联性, 其防御能力可以覆盖目前绝大多数拟态界内的未知漏洞、后门和病毒木马等安全威胁。尽管如此, DHR 构造在机制上可以自然地利用常规防御措施以增加拟态界内的相异性, 有助于非线性提高防御的有效性与可靠性。

6) 逆转攻防不对称格局

无论从定性定量分析还是实际测评[12]的角度, 都不难得出结论: DHR 的内在机理可以很容易获得非线性的防御增益[12], DHR 系统的失效率也是呈指数级衰减的。按照可靠性模型分析不难得出, 在相同余度数 $n(n \geq 3)$ 的条件下, DHR 的可靠性要远远高于非相似余度的可靠性; 从 DHR 的典型构造来看, 其防御成本上限正比于拟态界内异构执行体的数量 n, 为线性函数, 但其防御效果则是非线性增加的。更为重要的是, 从应用系统的全寿命周期来看, DHR 构造能显著降低实时防护性要求(如防御 0day 攻击等)、版本升级(打补丁)频度以及附加专门安全装置等所带来的维护成本; 从实现方面来看, DHR 允许使用全球化市场的 COTS 级软硬构件来组成异构执行体, 甚至可以直接使用开放或开源的产品。相比于专门设计、特殊制造的安全构件、部件或系统, DHR 组件的开发和售后服务成本可被规模化市场所消化, 基本上可以忽略。

7) 适应全球化生态环境

DHR 构造模式需要标准化、多样化、多元化的软硬构件市场和业态的支持, 而“开源社区”、全球化产业链等现代技术开发和生产组织方式恰好能起到自然的支撑作用。由于 DHR 构造的泛化使用能够提供更加强劲的多样化市场需求, 这也使得同质异

构技术产品将不只是排他性的竞争关系。

一般来说, 功能等价的软硬构件之间的技术成熟度必定存在差异, 特别是产品初期阶段或市场后来者。借助 DHR 构造的高可靠容错属性和多模判决的定位功能, 能快速发现拟态界内新构件产品的设计缺陷和性能弱点。预期可以实现以下三方面的效果: (1)用户能够不再为市场上缺乏成熟的多元化组件感到担忧和烦恼; (2)因为“同质异构”需求的出现, 可以大大降低新产品供应商进入市场的门槛; (3)包容设计缺陷的特性可以大大减少设计阶段系统验证和确认的时间及费用, 加快新品入市的进程。

DHR 虽然要求用功能等价的多元化或多样化软硬构件搭建运行环境, 但并不苛求拟态界内构件本身的“无毒无菌”或“绝对可信”。这使得我们可以用一些在安全性方面不完全可控但功能、性能和成熟度较高的国内外产品, 与自主可控程度较高但先进性或成熟性等方面尚存在差距的可信产品, 采取混合配置或伴随(可以是实时或准实时)监视的工作模式, 用功能性能和安全性可靠性等方面的高低搭配来优化应用系统全寿命周期内的成本结构或经济性指标。

8) 一体化的系统架构

显然, DHR 构造不仅具有很高的容错能力, 而且还能独立地提供确定或不确定的威胁防御能力。这种创新架构可以一体化地解决可靠、可信服务环境的鲁棒性、柔韧性和安全性的问题。相对于附加或堆砌在专门的安全防御设施来保护网络服务系统的方法, 其三大能力融为一体, 在部署上具有更佳的性价比或费效比。

4 网络空间拟态防御

4.1 目标与路线

网络空间拟态防御(Cyber Mimic Defense, CMD), 以下简称“拟态防御”, 旨在为解决网络空间不同领域相关应用层次上的基于未知漏洞、后门或病毒木马等不确定性威胁, 提供具有普适性的创新防御理论和方法。既能为关键网络设施或核心信息装备提供弹性化的或可重建的服务能力, 也能以一体化的架构技术提供独立于传统安全手段的内生安全增益, 或融合成熟的防御技术获得超非线性的防御效果。期望在网络空间营造一个与全球化时代技术和产业发展模式相适应, 合作共赢和开源众创相融合, 非封闭、自主可控、可持续发展的新兴生态环境。

CMD 在技术上以融合多种主动防御要素为宗旨: 以异构性、多样或多元性改变目标系统的相似性、单一性; 以动态性、随机性改变目标系统的静态性、确定性; 以异构冗余多模裁决机制识别和屏蔽未知缺陷与未明威胁; 以高可靠性架构增强目标系统服务功能的柔韧性或弹性; 以系统的视在不确定属性防御或拒止针对目标系统的不确定性威胁。用基于 DHR 的一体化技术架构集约化地实现上述目标。

4.2 基本概念

CMD 的基本概念可以简单归纳为“五个一”: 一个公理, “人人都存在这样或那样的缺点, 但极少出现在独立完成同样任务时, 多数人在同一个地方、同一时间、犯完全一样错误的情形”; 一种架构, 动态异构冗余 DHR 架构; 一种运行机制, “去协同化”条件下的多模裁决和多维动态重构机制; 一个思想, “移动攻击表面”(Moving Attack Surface, MAS)思想[13-15]; 一种非线性安全增益, 纯粹通过架构内生机理获得的拟态防御增益(Mimic Defense Gain, MDG)。其目的是在功能等价、开放的多元化生态环境上, 将复杂目标系统自主可控问题转化为功能相对单一的拟态括号之可控可信问题。

(1) 根据“给定功能和性能条件下, 往往存在多种实现算法”公理, 可证明这些实现方法的并集或交集运算结果仍然满足功能等价性要求。这意味着网络空间基于未知漏洞、后门等不确定威胁的防御难题, 能被异构冗余机制转化为可用概率描述的风险防护问题。

(2) 借助 MAS 思想, CMD 系统可以视为一种以攻击者不可预测的方式部署多维度攻击面的主动防御系统。由于各执行体的构造及环境存在时空维度上的多元异构性, 使得攻击者(包括潜伏者)可以利用的资源在时空维度上存在不确定性, 宏观上表现为攻击面总是在做不规则的移动。尤其对那些需要多步骤传送数据包或回送数据包才能达成目的的攻击任务, 其可达性前提几乎无法保证。

(3) “去协同化”条件下的多模策略裁决和多维动态重构机制, 能将复杂系统攻击表面的高难度高代价工程问题, 转变为空间独立的、功能简单的“拟态括号”之软硬部件攻击表面的缩小问题。并且, 能使异构冗余体之间可能存在的显性或隐性关联性降至最低, 给攻击者造成非配合条件下的异构多目标动态协同攻击困境。这使得自主可控的工程实现难度从全产业链“不得有安全短板”, 降低到只需在个别环节或关键部件“严防死守”即可。

(4) 拟态防御架构的安全增益 MDG 是“内生”的, 与现有的安全防护技术, 如加密认证、防火墙过滤、查毒杀毒、木马清除等入侵检测、预防、隔离

和清除措施, 在机理上无依赖关系, 漏洞修补、后门封堵或恶意代码清除等传统的增量修补手段只是作为稳定防御效果的补充性措施且无实时性要求。但是, 融合使用这些安全技术可使目标对象的防御能力获得超非线性提高。

4.3 拟态防御界

MDB 内包含若干组定义规范、协议严谨的服务(操作)功能。通过这些标准化协议或规范的一致性或符合性测试, 可判定多个异构(复杂度不限的)执行体在给定服务(操作)功能上甚至性能上的等价性。即通过拟态界面的输入输出关系的一致性测试可以研判功能执行体间的等价性, 包括给定的异常处理功能或性能的一致性。拟态界面所定义功能的完整性、有效性和安全性是拟态防御有效性的前提条件, 界面未明确定义的功能(操作)不属于拟态防御的范围(但也可能存在衍生的保护效应)。换句话说, 如果攻击行动未能使输出矢量不一致时, 拟态机制不会做出反应。因此, 合理设置、划分或选择拟态防御界在工程实现上就非常关键。

需要特别强调的是, 拟态界外的安全问题不属于拟态防御的范围。例如, 由钓鱼、在服务软件中捆绑恶意功能、在跨平台解释执行文件中推送木马病毒代码、通过用户下载行为携带有毒软件等不依赖拟态界内未知漏洞或后门等因素而引发的安全威胁, 拟态防御效果不确定。

4.4 拟态防御等级

(1) 完全屏蔽级

如果给定的拟态防御界内受到来自外部的入侵或“内鬼”的攻击, 所保护的功能、服务或信息未受到任何影响, 并且攻击者无法对攻击的有效性作出任何评估, 犹如落入“信息黑洞”, 称为完全屏蔽级, 属于拟态防御的最高级别。

(2) 不可维持级

给定的拟态防御界内如果受到来自内外部的攻击, 所保护的功能或信息可能会出现概率不确定、持续时间不确定的“先错后更正”或自愈情形。对攻击者来说, 即使达成突破也难以维持或保持攻击效果, 或者不能为后续攻击操作给出任何有意义的铺垫, 称为不可维持级。

(3) 难以重现级

给定的拟态防御界内如果受到来自内外部的攻击, 所保护的功能或信息可能会出现不超过 t 时段的“失控情形”, 但是重复这样的攻击却很难再现完全相同的情景。换句话说, 相对攻击者而言, 达成突破的攻击场景或经验不具备可继承性, 缺乏时间维度上可规划利用的价值, 称为难以重现级。

原则上可以定义更多的防御等级以适应不同应用场景对安全性与实现代价的综合需求。其中, 给攻击行动造成的不同程度的不确定性则是拟态防御的核心。不可感知性使得攻击者在攻击链的各个阶段都无法获得防御方的有效信息; 不可保持性使得攻击链失去可利用的稳定性; 不可再现性使得基于探测或攻击积淀的经验, 难以作为先验知识在后续攻击任务中加以利用等。

4.5 适用域

在具有函数化的输入输出关系或满足 IPO 模型的条件下, 用来解决服务的功能、性能和重要信息资源必须显性、健壮提供的问题; 对初始投资或产品价格不敏感; 空间或功耗条件可接受, 且存在标准或可归一化的功能接口与协议规范, 并具备多元或多样化处理条件; 对系统或装置的可靠性与可用性存在传统安全性和网络空间非传统安全双重要求的场合。

需要指出的是, 在满足上述条件下, 凡是目标算法确定且存在其他等价算法的应用场合都适合实施拟态防御。即使诸如科学计算类的课题或者工业控制领域中的应用, 由于功能等价的不同算法之间的结果精度可能存在差异(如精确到小数点后几位), 或者控制量的值域是一个范围, 也可以采用“精度掩码”或“预期范围”等策略的判定方式进行多模裁决

5 测试与评估[12]

5.1 背景情况

受国家科技部高新技术发展及产业化司委托, 上海市科委先后组织国内网络通信和安全领域一流科研院所、高等院校和知名企业的 21 名院士和 110 余名同行专家, 对拟态防御理论和验证系统进行测试评估。主要目的是测试“web 服务器拟态防御原理验证系统”、“路由器拟态防御原理验证系统”的安全功能、性能等是否达到设计预期, 为评估“拟态防御理论的正确性、有效性及工程实现的可行性”提供依据。

2016 年 1 月下旬至 4 月下旬为测试阶段。由来自国家信息技术安全研究中心、中国信息通信研究院、中科院信息工程研究所、军委装备发展部第六十一研究所、上海交通大学、浙江大学、北京奇虎科技有限公司、启明星辰信息安全技术有限公司、安天科技股份有限公司等 9 家单位优势测试团队的 45 名同行专家, 依据国家或行业相关标准、规范和拟态防御原理, 论证制定了测试验证方案, 并进行了三轮联合测试, 其中, 众测 103 项 194 例、互联网

渗透测试 10 项 10 例。

2016 年 5 月 7 日进入总结评估阶段。上海市科委召开了拟态防御原理验证系统测试情况评估会, 测试组汇报了测试工作组织实施情况和测试结论, CMD 联合研发团队做了测试数据分析报告。2016 年 5 月 29 日, 科技部高新司、中国工程院信息与电子工程学部和上海市科委, 在北京联合召开了拟态防御原理暨应用实践研讨会。参与测试评估工作的同行专家以通信方式经过充分评议和修改, 于 2016 年 8 月正式形成《拟态防御原理验证系统测评意见》。

5.2 测试概况

联合测试团队, 在确认提交的受测系统本征功能和性能满足国家或行业标准且测试验证全过程可保持的前提下, 通过拟态及非拟态模式对比方式验证 CMD 防御的有效性, 通过渗透测试验证 CMD 对各种渗透攻击的防御能力, 通过“白盒”及“配合植入”后门或病毒木马等开放式测试手段, 检验 CMD 构造在“去协同化”条件下的协同攻击难度。

受测对象之一, web 服务器拟态防御原理验证系统, 共完成 7 类 70 项 161 例测试验证。内容包括功能测试、HTTP1.1 协议一致性测试、安全性测试、接入测试、性能测试、兼容一致性测试和互联网渗透等测试验证。

受测对象之二, 路由器拟态防御原理验证系统, 共完成 6 类 43 项 43 例测试验证。内容包括安全性测试、OSPF 协议功能测试、性能测试和互联网渗透等测试验证。

5.3 初步结论

受测系统是拟态防御理论与方法的成功实践。在满足通用 web 服务器和专用路由器功能、性能标准要求的同时, 实现了基于异构冗余多维动态重构的主动防御机制, 即 CMD 机制, 能够独立且有效地应对或抵御 MDB 内基于漏洞、后门等的已知风险或不确定威胁, 其叠加与迭代效应能够非线性地增加目标对象的攻击难度。已作的测试验证表明, 受测系统达到拟态防御理论预期。

“灰盒和白盒”测试验证结果表明, 现有的扫描探测、漏洞利用、后门设置、病毒注入、木马植入乃至 APT 等攻击手段和方法, 对 MDB 内受保护对象没有预期的作用和可信效力。

综合测试分析表明, 在功能等价异构冗余的多维动态重构机制作用下, MDB 内几乎不可能实现可靠、持续的协同逃逸。同时, CMD 机制还具有大幅度降低系统全寿命周期内专用安全设施的更新升级代价、防护实时性要求、版本同步更新频度等综合优势。在产业链开放的全球化生态环境中, 使得利用“有毒带菌”构件实现可管可控的信息系统成为可能, 给基于“后门工程和隐匿漏洞”等攻势战略带来颠覆性影响。

测试验证还表明, 拟态防御不仅具有显著的安全功效, 同时还能够提供期望的服务功能与高可靠的应用场景, 并能自然地继承和接纳网络安全与信息化领域的科技成果。普适性的系统架构开辟了内生防御理论和技术研究新方向, 基于“不对称”优势为网络空间“自主可控”战略探索出克服瓶颈问题的新途径, 具有“改变游戏规则”的重要意义。

理论预期和测试验证显示, 非 MDB 内的安全问题, 如基于网络协议、服务功能算法等设计缺陷, 或利用社会工程学方法和手段等实施的攻击, CMD 亦有程度不同的防护效果。

6 总结与展望

CMD 为用系统架构技术解决了 MDB 内基于未知漏洞、后门或病毒木马等不确定威胁, 提供了“改变游戏规则”的新途径。CMD 的内生安全增益既能独立于传统的安全防御手段, 也能很好地综合后者的优势; 其开放性与安全性的完美结合将促进全球化进程向深度广度发展; 其系统构造具有普适性, 能够集约化地实现网络服务、可靠性保障与安全防御等功能。

参考文献

[1] Knight, Hyneman F. “Risk, uncertainty and profit.” *Houghton Mifflin Company*, 1921.

[2] Kocher P C. “Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems.” *Advances in Cryptology*. pp. 104-113, 1999.

[3] Liu, Fangfei, et al. “Last-level cache side-channel attacks are practical.” In *Security and Privacy (S&P’15)*, pp. 605-622, 2015.

[4] “Mimic octopus”. Wikipedia, the Free Encyclopedia. https://en.wikipedia.org/wiki/Mimic_octopus.

[5] Voas J, Ghosh A, Charron F, et al. “Reducing uncertainty about common-mode failures.” In *Proc. IEEE Symp. Software Reliability Engineering (SRE’97),* pp. 308-319, 1997.

[6] Levitin G. “Optimal structure of fault-tolerant software systems.” *Reliability Engineering & System Safety*, vol. 89, no. 3, pp. 286-295, 2005.

[7] Yeh Y C B. “Triple-triple redundant 777 primary flight computer.” In *Proc. Aerospace Applications Conference (AAC’ 96)*, pp. 293-307, 1996.

[8] Denning D E. “An intrusion-detection model.” *IEEE Transactions on Software Engineering,* vol. 13, no. 2, pp. 222-232, 1987.

[9] Kreibich C, Crowcroft J. “Honeycomb: creating intrusion detection

signatures using honeypots." *ACM Sigcomm Computer Communication Review,* vol. 34, no. 1, pp. 51-56, 2015.

[10] Pal P, Webber F, Schantz R E, et al. "Intrusion tolerant systems," In *Proc. IEEE. Information Survivability Workshop (ISW-2000).* pp. 24-26, 2000.

[11] Nguyen Q L, Sood A. "A comparison of intrusion-tolerant system architectures." *Security & Privacy IEEE*, vol. 9, no. 4, pp. 24-31, 2011.

[12] 上海市科学技术委员会给国家科技部高新技术发展及产业化司的专题报告, "拟态防御原理验证系统测试评估工作情况汇总", 2016. 8

[13] Jajodia S, Ghosh A K, Swarup V, et al. "Moving target defense." *Springer*, 2011.

[14] Manadhata P K, Wing J M. "An attack surface metric." *Software Engineering IEEE Transactions on,* vol. 37, no. 3, pp. 371-386, 2011.

[15] Manadhata P K. "Game theoretic approaches to attack surface shifting." Moving Target Defense II. *Springer*, 2013.

邬江兴, 现在国家数字交换系统工程技术研究中心主任, 教授, 博导。研究领域为信息通信网络、网络安全。Email: 17034203@qq.com