

基于 GSCPn 的硬件木马攻击建模与安全性分析

梅波, 郭威, 曹志鹏, 李沛杰, 张文博
(信息工程大学, 河南 郑州 450001)

摘要: 针对硬件木马 (hardware trojan, HT) 的研究主要集中在 HT 设计和防御的具体策略, 缺乏面向 HT 攻击的系统安全性量化评估模型的问题, 提出了基于广义随机着色 petri 网 (generalized stochastic coloring petri net, GSCPn) 的 HT 攻击和安全评估模型。基于单个基本单元的 HT 攻击行为和状态之间的时序关系, 构建了单个基本单元的 HT 攻击 GSCPn 模型, 根据系统工作状态下基本单元的通信依赖关系, 给出了构建系统 HT 攻击 GSCPn 模型的通用方法。求解了与 GSCPn 模型同构的马尔可夫链的解析解, 以及采用蒙特卡罗模拟方法的仿真解, 二者的误差仅为 0.001 83, 验证了所提方法建模的准确性。通过案例分析了不同场景下系统的安全性, 实验结果表明, 木马攻击强度、木马恢复时间和任务执行时间均对安全性有不同程度的影响。

关键词: 硬件木马; 广义随机着色 petri 网; 攻击建模; 连续时间马尔可夫链

中图分类号: TP309

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024054

Modeling and security analysis of hardware trojan attack based on GSCPn

MEI Bo, GUO Wei, CAO Zhipeng, LI Peijie, ZHANG Wenbo
Information Engineering University, Zhengzhou 450001, China

Abstract: Addressing the issue that research on hardware Trojan (HT) has predominantly focused on specific strategies of HT design and defense, with a deficiency in quantitative evaluation models for system security in the context of HT attacks, a HT attack and security assessment model based on the generalized stochastic coloring Petri net (GSCPn) was proposed. The GSCPn model representing the HT attack of a single basic unit was constructed, founded on the temporal relationship between HT attack behaviors and the states of a single basic unit. A general method for constructing the GSCPn model of HT attack for the entire system was provided, based on the communication dependency relationships between the basic units during the system's operational state. The analytical solution of the Markov chain isomorphous to the GSCPn model was derived, alongside the simulation solution

收稿日期: 2024-03-10; 修回日期: 2024-08-09

通信作者: 郭威, guowjss@126.com

基金项目: 国家重点研发计划 (2022YFB4401401)

Foundation Item: The National Key Research and Development Program (2022YFB4401401)

引用格式: 梅波, 郭威, 曹志鹏, 等. 基于 GSCPn 的硬件木马攻击建模与安全性分析[J]. 网络与信息安全学报, 2024, 10(4): 72-84.

Citation Format: MEI B, GUO W, CAO Z P, et al. Modeling and security analysis of hardware trojan attack based on GSCPn [J]. Chinese Journal of Network and Information Security, 2024, 10(4): 72-84.

achieved using the Monte Carlo simulation method. The discrepancy between the two methods was a mere 0.00183, which verified the precision of the modeling approach. The security of the system under various scenarios was analyzed through a straightforward case study. The experimental results indicate that the intensity of the HT attack, HT recovery time, and task execution time all exert varying degrees of influence on system security.

Keywords: hardware trojan, generalized stochastic coloring petri net, attack modeling, continuous time Markov chain

0 引言

5G基站、大数据中心等新型基础设施建设(新基建)与安全关键型领域(如医疗、交通)深度融合。以及新基建的发展离不开底层硬件平台的支持,然而半导体行业发展的全球化生态,多个不可信第三方的加入给硬件平台带来严峻的硬件安全挑战,如硬件木马(hardware trojan, HT)和盗版问题,其中HT问题是近年来研究的热点问题。HT在2007年被首次提出^[1],在过去的十几年里,学者已经在HT领域开展了大量的研究。主流的HT研究可以分为两大分支:木马攻击研究和木马防御研究。

为了成功实施HT攻击,HT攻击主要围绕增加木马植入概率、增强木马隐蔽性、木马触发机制和木马负载机制4个方面进行研究。Chen等^[2]分析了分割制造方法的弊端,提出了一种基于接近度的攻击方法,该方法可以恢复部分连接信息,从而便于植入木马。Khan等^[3]提出了一种木马的负载逻辑和触发逻辑位于不同单元的分布式协同木马攻击场景。Hu等^[4]提出了一种位于深度神经网络加速器中的HT,该木马的触发条件为输入图像数据的特定几何特征,恶意负载在木马触发后施行目标或非目标攻击,这种木马因容易被攻击者精确控制而更具有威胁性。

HT的防御方法可以分为木马检测方法、可信设计方法和运行时防护的方法。木马检测的方法旨在集成电路的各个生命周期,实现对木马的检测。Yasaei等^[5]将电路的HDL代码转换成特定的数据流图,然后利用图卷积网络对图中节点进行分类,实现硅前的木马检测与定位。Yang等^[6]通过采集逻辑测试时产生的侧信道信息对硬件进行聚类,然后从每个类中选择一个硬件进行逆向工程,实现硅后的木马检测。可信设计技术可以通过改变设计制造流程,以此增加木马植入的难度,如分割制造^[7]和布局混淆^[8],也可以通过植

入安全模块如传感器^[9]增强木马的检测准确度。木马检测和可信设计技术仍无法完全保证系统无HT,运行时防护方法作为HT防御的最后一道防线,旨在提高系统运行时的安全性。Dong等^[10]设计了一个多级架构,通过引入安全的包装器和控制器实现对第三方加密IP木马的检测与保护。Cassano等^[11]提出一种软件混淆方法DETTON,通过生成受保护软件的混淆版本减轻被窃取信息木马攻击的危害。

HT攻击和防御方法促进了HT领域的发展,但是目前的研究大多专注于具体的攻防技术细节,缺乏一个可以建模HT攻击的系统安全评估方法。通过构建硬件木马攻击模型,在被攻击前进行硬件安全风险评估,并量化识别模型中安全关键点,采取合适安全防御方法,以低防御成本实现有效的木马容忍与防御。同时也可以模型中扩展防御行为,从系统的视角横向对比不同防御方法的安全性。

基于上述分析,为了弥补在HT攻击建模与安全性分析研究方向的空白,从系统层面分析HT攻击对系统安全性的影响,以系统视角安全性提升为目标指导防御者采取合适的防御措施,本文提出了基于广义随机着色petri网的分布式协同HT攻击模型。本文的主要工作如下。

1) 本文分析了一般木马和分布式协同木马的攻击流程,基于木马的恶意功能利用广义随机着色petri网(generalized stochastic coloring petri net, GSCP)对系统基本单元的HT攻击进行建模,并根据系统模型给出了构建完整系统HT攻击GSCP模型的通用方法。

2) 本文将与基本单元GSCP模型同构的连续时间马尔可夫链(continuous time Markov chain, CTMC)稳态分布解析解和基于蒙特卡洛模拟方法的仿真解进行对比,二者的误差仅为0.001 83,验证了本文建模方法的准确性。

3) 通过案例分析了不同场景下系统的安全

性指标, 实验结果表明, HT 攻击强度、HT 恢复时间、任务执行时间和基本单元的不同逻辑位置对系统的安全性有不同程度的影响。

1 硬件木马与网络攻击模型

1.1 硬件木马

HT 是能够被攻击者利用的特殊硬件模块, 由触发逻辑和负载逻辑组成, 触发逻辑可以进一步分为组合逻辑和时序逻辑触发, 当满足触发条件时, 激活负载逻辑, 从而实现特定的恶意功能以达到攻击者意图^[12]。硬件木马逻辑如图 1 所示。除了依据触发逻辑的分类, HT 还可以依据其他不同的分类标准进行分类, 如图 2 所示。

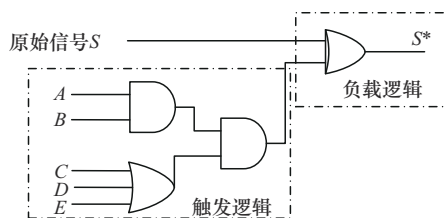


图1 硬件木马逻辑
Figure 1 HT logic

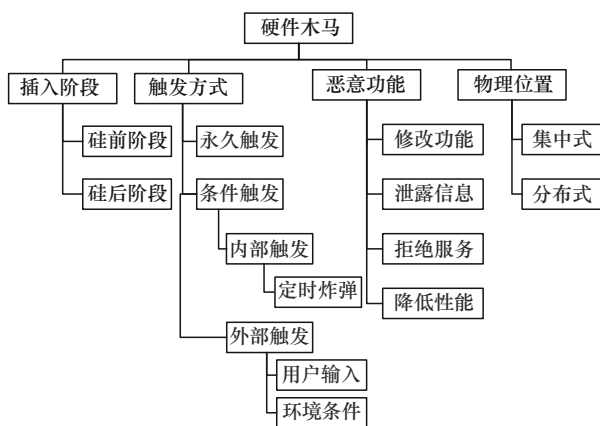


图2 硬件木马分类
Figure 2 HT classification

随着分布式异构多核系统的发展, 基于物理位置分类的分布式 HT^[13-16] 攻击场景也受到研究者的关注。如片上多处理器系统 (multi-processor system on chip, MPSOC) 中的第三方 IP 供应商在其提供的不同 IP 中分别植入 HT 的触发和负载逻辑, 基于 Chiplet 的系统会受到来自 Chiplet 中 IP 供应商、不可信 Chiplet 供应商、不

可信制造商的分布式 HT 攻击 (此处的攻击是指植入分布式 HT)。此类 HT 由于更为隐蔽, 容易逃过系统运行前的各种检测而更受攻击者青睐。分布式 HT 的完整激活需要两个条件: ①具有满足触发逻辑模块中的触发条件。②满足触发条件后, 具有触发逻辑的模块通过与具有负载逻辑的模块交互, 隐蔽地将 HT 激活, 使信号能传递至负载逻辑, 实现木马激活。由上述分析可知, 条件①和条件②具有严格的时序关系, 只有条件①先于条件②被满足时, HT 才会完整被激活。

1.2 网络攻击模型

HT 攻击是一种特殊的网络攻击形式, 因此可以参考已有的网络攻击模型来建立 HT 攻击模型。学术界提出了多种攻击模型, 并基于攻击模型进行网络安全风险评估, 目前较典型的攻击模型主要有攻击树模型、攻击图模型以及攻击网模型。

1.2.1 攻击树模型

攻击树模型^[17-19]由可靠性领域的故障树模型演化而来, 它通过自下而上的分析方法, 刻画网络攻击。模型由叶节点、子节点和根节点组成。叶节点表示攻击的基本操作, 根节点表示攻击的最终目标, 由叶节点到根节点的一条完整路径可表示实现攻击目标的一个完整攻击过程。攻击树的建模方法可以从攻击者视角表示开始实施到完成最终目标的攻击过程, 但是攻击树模型不能描述攻击行为和系统状态的动态关系, 缺乏表达并发攻击和协同攻击的能力。

1.2.2 攻击图模型

Phillips 等^[20]提出了攻击图的建模方法。攻击图是一种有向图, 节点之间通过有向边连接。攻击图可以细化为状态攻击图和属性攻击图^[21]。在状态攻击图中, 节点表示系统的所处的状态信息, 边表示攻击行为如漏洞利用。属性攻击图具有两类节点, 条件节点和脆弱性节点, 只有当脆弱性节点的所有父条件节点都被满足时, 攻击者才可以利用该脆弱性进行攻击, 脆弱性节点的所有子节点表示攻击者利用该脆弱性进行攻击之后系统的状态。攻击图模型可以展示多个系统脆弱性之间的利用关系, 并通过对有向边赋予时间权重表示时间属性, 但是攻击图模型同样存在不能很好地对并发和协同攻击建模的问题。

1.2.3 攻击网模型

攻击网模型是一类利用petri网对攻击进行建模的技术，petri网的基本元素包括库所和变迁，库所表示条件或状态，变迁表示事件，petri网最初被广泛应用于具有分布式、并发等属性的系统建模。petri网具有出色的图形表达能力，能够同时对攻击行为和系统运行状态进行刻画，并反映二者之间的时序关系。petri网会遇到状态空间爆炸的问题，研究人员提出将petri网扩展形式如着色petri网^[22]、随机petri网^[23]和层次petri网^[24]等应用于网络攻击建模、可靠性和安全性分析的方法。

由1.1节可知，HT的激活是一个动态的过程，木马攻击进行到不同阶段，系统会处于不同的状态，且分布式异构多核系统中存在多个核同时被HT攻击的并发现象。为了清晰地表述HT的攻击过程和系统状态的因果和时序关系，并对并发HT攻击和分布式HT攻击进行建模，同时缓解状态空间爆炸问题，本文采用GSCPN对HT攻击进行建模。

2 硬件木马攻击GSCPN模型

2.1 威胁假设

没有任何一个攻击模型可以适用于所有的系统，也无法囊括所有的攻击。因此，在建立攻击模型之前，先阐述本文研究的系统模型和威胁模型，声明系统边界和威胁边界，所提出的攻击模型只适用于边界内部。

2.1.1 系统模型

(1) 系统组成

本文所述系统是一种异构多核系统，由若干(≥ 1)可能被植入HT的模块(核)组成，称这种模块为基本单元。模块可以有不同粒度的定义，在本文中，模块由系统集成时最直接的不可信任第三方提供，如基于Chiplet集成的系统，以单个Chiplet为基本单元，而不以单个Chiplet中集成的第三方IP核为基本单元。

(2) 系统功能

系统的功能是执行用户提交的应用，应用由一组具有通信关系的子任务组成，用有向无环图(directed acyclic graph, DAG)对应用进行建模，本文定义表示应用的DAG图 $G=(V,E), V=$

$\{v_1, v_2 \cdots v_p\}$ 为任务节点集合， E 为有向边集合， $e_{ab} = (v_a \rightarrow v_b) \in E$ ，则说明任务 a 与任务 b 之间有通信依赖关系，即 b 的执行依赖 a 的执行结果，称任务 a 为任务 b 的一个父任务，同时将执行任务 a 的基本单元称为执行任务 b 的基本单元的父单元。

2.1.2 威胁模型

(1) 触发机制

HT只在基本单元工作时概率触发。基本单元工作时到达的特殊系统状态，如特殊指令序列、数据输入或者温度电压等物理条件均可触发HT。当达到木马触发条件时称为木马触发(但不一定立刻被激活)。

(2) 负载机制

假设不可信第三方提供的基本单元中可以同时存在泄露信息、篡改功能和拒绝服务3种不同恶意负载的HT。其中篡改功能木马会对应用的输出有修改，拒绝服务木马导致基本单元发生不可逆损坏，本文称木马执行恶意负载功能为木马激活。

(3) 恢复休眠机制

除拒绝服务木马外，已激活的木马会在触发条件由被满足转向不被满足时自动恢复至休眠状态，或在基本单元任务执行完成时被动恢复至休眠状态。

(4) HT物理位置

同时存在集中式和分布式HT，分布式HT位于同一第三方提供的不同基本单元中。当分布式HT触发逻辑的基本单元为负载逻辑基本单元的父单元，二者具有合法的通信交互行为时，HT才会被概率激活，否则分布式HT无法被激活，本文不考虑分布式HT利用热隐蔽信道等非合法通信信道交互协同激活HT的情况。

2.2 广义随机着色Petri网攻击模型

为了便于后续利用连续时间马尔可夫链对所提攻击模型进行量化分析，本文假设HT的攻击扰动到达时间、基本单元执行任务的截止时间、HT自动恢复时间均是无记忆的，服从负指数分布。最后，为了简化后续安全性分析，假设同一时刻单个基本单元中最多只有一个HT处于激活状态，其他HT处于休眠状态。

基于GSCPN的基本单元HT攻击模型可描述九元组为GSCPN-BUHT = $\{\Sigma, P, T, F, C, G, \omega, \lambda, M_0\}$ 。其中： Σ 为网模型中的颜色集，由基本色和混合

色组成, 本文图中用<>符号及其内部参数表示颜色, GSCPN-BUHT 的颜色集及其含义如表 1 所示。 P 为网模型中库所的集合, 图中用圆形表示。 T 为网模型中变迁的集合, $T = T_{\text{time}} \cup T_{\text{immediate}}$, $T_{\text{time}} \cap T_{\text{immediate}} = \emptyset$, T_{time} 为时延变迁集合, $T_{\text{immediate}}$ 为瞬时变迁集合, 图中分别用空心矩形和实心矩形表示。 F 为网模型中的有向弧集合, $F \subseteq (P \times T) \cup (T \times P)$, 弧只可以由变迁指向库所或者库所指向变迁, 图中由带有箭头的连线表示。 C 为颜色映射函数, $C: P \rightarrow 2^{\Sigma}$, 2^{Σ} 为颜色集的子集, 表示为库所可能的令牌色的集合。 G 为网模型中变迁的守卫函数集, $G: T \rightarrow \text{Bool Expression}$, 当满足表达式条件时变迁才可能被激活, 无特殊说明情况下 $\text{Bool Expression} = 1$ 。 ω 为瞬时变迁的归一化权重, 满足 $\sum_{t \in T_{\text{immediate}}} \omega_t = 1$, 表示在标识 M 下, 所有可实施的瞬时变迁的权重和为 1, 无特殊说明情况下, 某一标识下的瞬时变迁发生概率相同即权重相同。 λ 为时延变迁的平均实施速率。 M_0 为模型初始标识。

一个基本单元 i 的 HT 攻击 GSCPN 模型的图形化描述如图 3 所示。在图 3 中, 攻击模型由 3 个部分组成, 说明如下。

(1) Part1

基本单元的父单元对其初始运行状态的影响, 由父单元分布式协同木马是否触发 (库所 $P_i_parent_out$, $* \in [1, h]$ 及其令牌颜色) 和与父单元是否来自同一第三方 ($P_i_parent_vendor$ 及其令牌颜色) 决定。影响逻辑由守卫函数

$\text{Guard}(T_i_parent_infl)$ 定义。变迁 $T_i_parent_all$ 表示所有父单元对基本单元状态的影响。影响逻辑由守卫函数 $\text{Guard}(T_i_parent_all)$ 定义。模型守卫函数及其表达式如表 2 所示。

(2) Part2

基本单元在运行时其状态的变化。库所 P_i 含有的混合色令牌表示基本单元的状态, 变迁 T_i_att 表示基本单元中木马触发的事件, 服从参数 λ_{att_i} 为的负指数分布, 变迁 T_i_rec 表示基本单元中已激活木马主动恢复事件, 服从参数为 λ_{rec_i} 的负指数分布, 变迁 T_i_task 表示基本单元完成任务的事件, 服从参数为 λ_{task_i} 的负指数分布。变迁 $T_i_task_disrupt$ 表示拒绝服务木马激活导致任务执行失败。变迁 $T_i_att_scenario$ 表示木马触发条件到达之后的具体攻击行为。库所 $P_i_task_begin$ 含有令牌表示基本单元在执行任务, $P_i_task_end$ 含有令牌表示基本单元执行任务结束。库所 P_i_broken 含有令牌表示基本单元拒绝服务木马攻击导致永久失效。库所 $P_i_att_before$ 含有令牌表示基本单元在木马触发时刻可能的状态。

(3) Part3

库所 $P_i_out_n$ 及其着色令牌表示基本单元在执行完任务时, 分布式协同木马是否触发, 库所将是否会作为基本单元的后续叶子节点单元木马攻击模型中 part1 部分的父节点输出库所。

基于基本单元的 GSCPN-BUHT, 可以构建系统的木马攻击模型, 方法如下。

Input: DAG of Application、GSCPN-BUHT

Output: GSCPN-DHMHT

表 1 颜色集及其含义
Table 1 Color sets and meanings

颜色子集	令牌颜色	含义
Cooperate	<yes>	满足分布式木马触发条件
	<no>	不满足分布式木马触发条件
HT	<nor>	基本单元无木马被激活
	<mod>	基本单元篡改功能木马被激活
	<lek>	基本单元信息泄露木马被激活
	<dos>	基本单元拒绝服务木马被激活
Vendor	<sam>	父单元与基本单元来自同一第三方
	<dif>	父单元与基本单元来自不同第三方
Basic unit	<Cooperate, HT>	基本单元的状态: 由分布式协同木马触发逻辑是否触发、木马是否激活共同决定

基本色	Cooperate	=	{yes,no}	颜色变量	j,k	:	Cooperate
基本色	HT	=	{nor,mod,lek,dos}	颜色变量	x,y,z	:	HT
基本色	Vendor	=	{sam,dif}	颜色变量	m	:	Vendor
混合色	Basic unit	=	Cooperate*HT	参数变量	h,n	:	表示父节点个数和子节点个数
				参数变量	i	:	表示具体基本单元

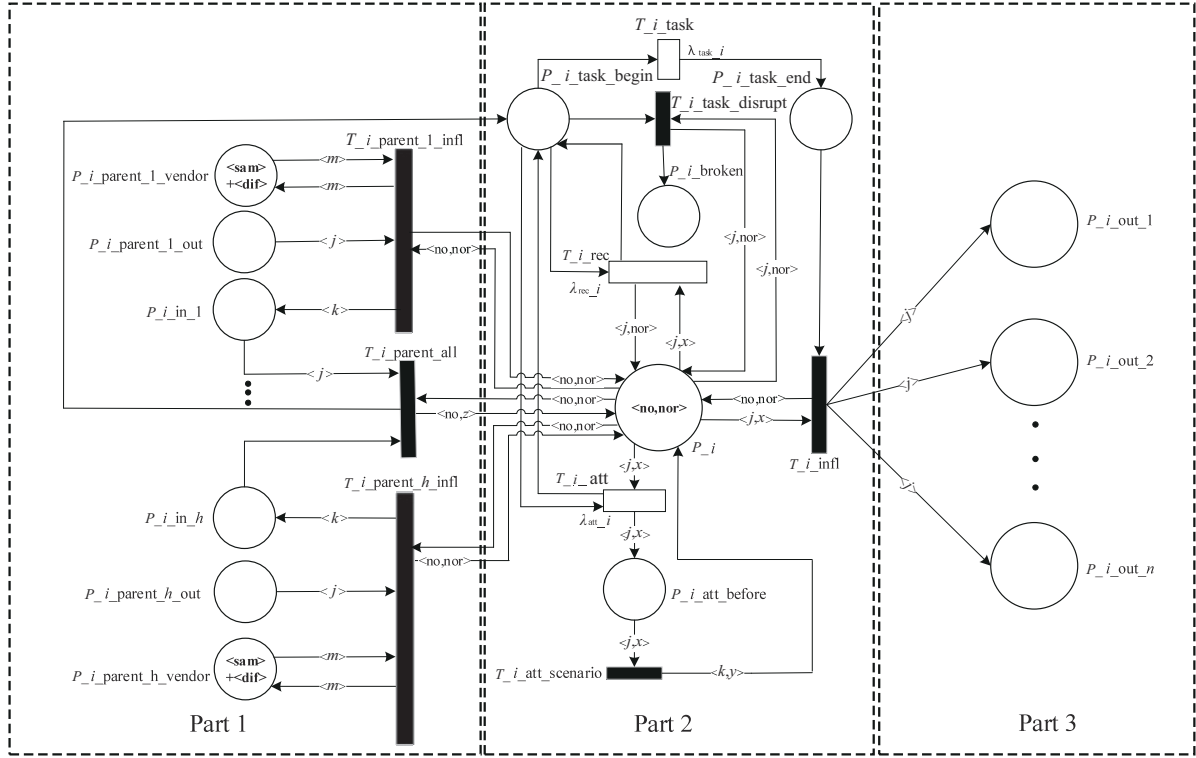


图3 基本单元HT攻击GSCP模型

Figure 3 GSCP model of basic unit HT attack

表2 模型守卫函数及其表达式

Table 2 Model guard functions and their expressions

函数	布尔表达式
Guard($T_{i_parent_*_infl}$)	$((j = \text{nor} \vee m = \text{dif}) \wedge k = \text{no}) \vee (j = \text{yes} \vee m = \text{sam} \wedge k = \text{yes})$
Guard($T_{i_parent_all}$)	$(j = \text{no} \wedge \dots \wedge k = \text{no} \wedge z = \text{nor}) \vee ((j \neq \text{no} \vee \dots \wedge k \neq \text{no}) \wedge z \neq \text{nor})$
Guard(T_{i_att})	$(j = \text{yes} \wedge x = \text{nor}) \vee (j \neq \text{yes} \wedge x \neq \text{dos})$
Guard(T_{i_rec})	$x \neq \text{nor} \wedge x \neq \text{dos}$
Guard($T_{i_att_scenario}$)	$(x \neq \text{nor} \wedge j = \text{no} \wedge k = \text{yes} \wedge y = x) \vee (x = \text{nor} \wedge j = \text{yes} \wedge k = j \wedge y \neq \text{nor}) \vee (x = \text{nor} \wedge j = \text{no} \wedge (k = \text{yes} \vee y \neq \text{nor}))$

Begin:
models_list = []
for each v_a in V : //生成每个基本单元的HT攻击模型
GSCP-BU_a_HT ← copy of GSCP-BUHT
calculate In-degree(v_a), Out-degree(v_a)
revise(GSCP-BU_a_HT): //对part1 和part3 进行修改

variable parameter i in GSCP-BU_a_HT ← a
 $P_{a_parent_out}$ = set()
 P_{a_out} = set()
if In-degree(v_a) = 0:
delete part1 of GSCP-BU_a_HT
 $P_{a_task_begin}$ ← one token
else:
parameter h in GSCP-BU_a_HT ← In-

```

degree( $v_a$ )
  for  $q$  in  $(1, h)$  :
     $P\_a\_parent\_out.append(P\_a\_parent\_q\_out)$ 
    if  $Out-degree(v_a) = 0$ :
      delete part3 of GSCPN-BU_ $a\_HT$ 
    else:
      parameter  $n$  in GSCPN-BU_ $a\_HT \leftarrow Out-$ 
degree( $v_a$ )
  for  $g$  in  $(1, n)$  :
     $P\_a\_out.append(P\_a\_out\_g)$ 
    models_list.append(GSCPN_BU_ $a\_HT$ )
for each  $e_{ab}$  in  $E$ : //依据节点间连接关系将基本单
元 HT 攻击模型组合
   $P\_a\_out\_s \leftarrow get\_random\_place(P\_a\_out)$ 
   $P\_b\_parent\_u\_out \leftarrow get\_random\_place$ 
( $P\_b\_parent\_out$ )
   $P\_a\_out\_s = P\_b\_parent\_u\_out$ 
   $P\_a\_out.delete(P\_a\_out\_s)$ 
   $P\_b\_parent\_out.delete(P\_b\_parent\_u\_out)$ 
End

```

3 模型正确性验证

本节假设系统只有一个基本单元，通过建模单个基本单元执行一次任务的 HT 攻击 GSCPN 模型，然后基于模型求得系统状态分布的解析解，并将其与基于蒙特卡洛模型求得的系统状态分布仿真解进行对比。所述简单系统的 GSCPN 模型和等价广义随机 petri 网 (generalized sochastic petri net, GSPN) 模型如图 4 和图 5 所示。

GSPN 模型可达标识集如表 3 所示，GSPN 的可达情态集分为实存态 T 和消失态 V ，与其同构的 CTMC 对应的转移概率函数矩阵为：

$$U = A + B = \begin{bmatrix} C & D \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ E & F \end{bmatrix} \quad (1)$$

其中， C 、 D 、 E 、 F 分别为 V 转换至 V ， V 转换至 T ， T 转换至 V ， T 转换至 T 的转移概率函数矩阵。

将 CTMC 中的消失态移除，得到压缩连续时间马尔可夫链 (reduced continuous time Markov chain, RCTMC) 的转移概率函数矩阵：

$$U' = [u_{ij}] \quad (2)$$

$$u_{ij} = f_{ij} + \sum_{r \in V} P_r[r \rightarrow j] \quad (3)$$

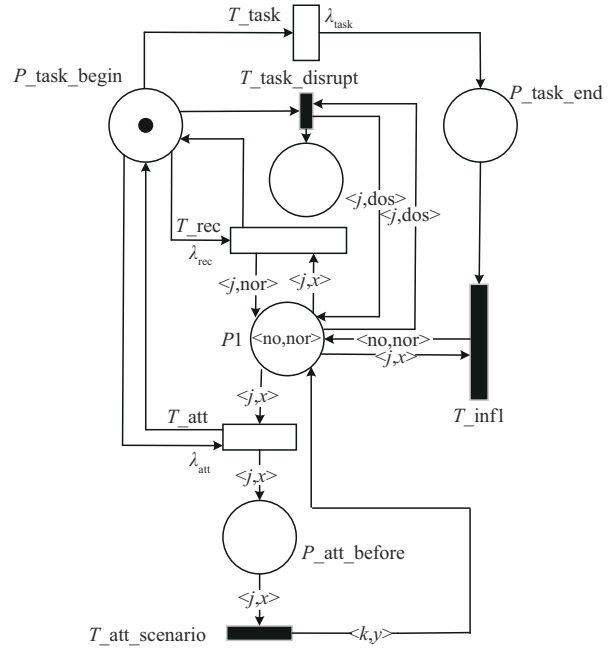


图 4 GSCPN 模型

Figure 4 GSCPN model

其中 $i, j \in T$, $f_{ij} \in F$, $P_r[r \rightarrow j]$ 表示由状态 $r \in V$ 转移任意步数到状态 j 的概率，转移路径中无实存态， $P_r[r \rightarrow j]$ 由矩阵 G^∞ [25] 得到：

$$G^\infty = [g_{rj}] = \begin{cases} \left(\sum_{h=0}^{k_0} C^h \right) D, & \text{消失态无循环} \\ [I - C]^{-1} D, & \text{消失态有循环} \end{cases} \quad (4)$$

令 $P_r[r \rightarrow j] = g_{rj}$ ，则对 $\forall i, j \in T, r \in V$ ：

$$u_{ij} = f_{ij} + \sum_{r \in V} P_r[r \rightarrow j] = f_{ij} + \sum_{r \in V} e_{ir} g_{rj} \quad (5)$$

$$U' = F + E G^\infty \quad (6)$$

由转移概率函数矩阵可得转移速率矩阵 Q ，其中，转移速率矩阵元素 q_{ij} 为：

$$q_{ij} = \begin{cases} \lim_{\Delta t \rightarrow 0} \frac{u_{ij}(\Delta)t}{\Delta t}, & i \neq j \\ \lim_{\Delta t \rightarrow 0} \frac{u_{ij}(\Delta t) - 1}{\Delta t}, & i = j \end{cases} \quad (7)$$

由柯尔莫哥洛夫向前方程可得：

$$\begin{cases} P'(t) = P(t)Q \\ P(0) = (P_1(0), P_2(0), \dots, P_n(0)) \end{cases} \quad (8)$$

且实存态的稳态概率满足：

$$\begin{cases} P \times Q = 0 \\ \sum_{i=1}^n P_i = 1 \end{cases} \quad (9)$$

在本节所设模型中， Q 矩阵如下所示：

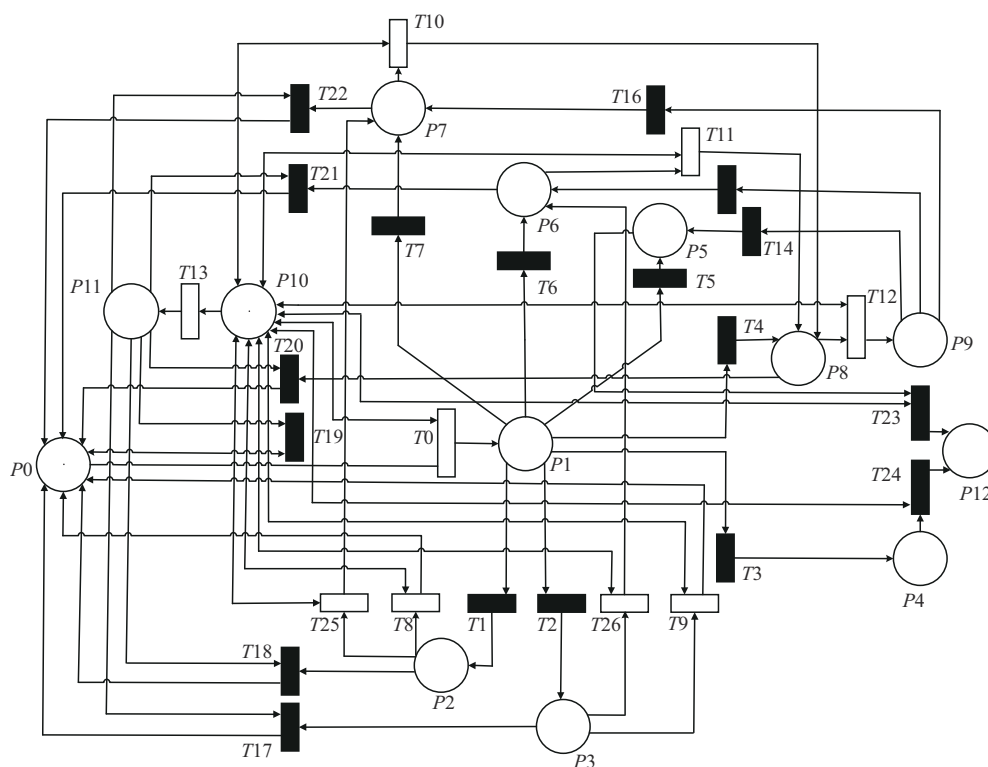


图5 等价GSPN模型

Figure 5 Equivalent GSPN model

表3 GSPN模型可达标识集及库所对应令牌

Table 3 GSPN reachable identification sets and tokens corresponding to places

	P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	备注
M_0	1	0	0	0	0	0	0	0	0	0	1	0	0	实存态
M_1	0	0	1	0	0	0	0	0	0	0	1	0	0	
M_2	0	0	0	1	0	0	0	0	0	0	1	0	0	
M_3	0	0	0	0	0	0	0	0	0	0	0	0	1	
M_4	0	0	0	0	0	0	1	0	0	0	1	0	0	
M_5	0	0	0	0	0	0	0	1	0	0	1	0	0	
M_6	0	0	0	0	0	0	0	0	1	0	1	0	0	
M_7	0	0	0	0	0	0	0	0	0	0	0	0	1	消失态
M_8	0	1	0	0	0	0	0	0	0	0	1	0	0	
M_9	0	0	0	0	0	0	0	0	0	1	1	0	0	
M_{10}	0	0	0	0	1	0	0	0	0	0	1	0	0	
M_{11}	0	0	0	0	0	1	0	0	0	0	1	0	0	
M_{12}	0	0	1	0	0	0	0	0	0	0	0	1	0	
M_{13}	0	0	0	1	0	0	0	0	0	0	0	1	0	
M_{14}	0	0	0	0	0	0	1	0	0	0	0	1	0	
M_{15}	0	0	0	0	0	0	0	1	0	0	0	1	0	
M_{16}	0	0	0	0	0	0	0	0	1	0	0	1	0	
M_{17}	1	0	0	0	0	0	0	0	0	0	0	1	0	

$$Q = \begin{bmatrix} -(\lambda_{att} + \lambda_{task}) & \frac{\lambda_{att}}{7} & \frac{\lambda_{att}}{7} & \frac{2\lambda_{att}}{7} & \frac{\lambda_{att}}{7} & \frac{\lambda_{att}}{7} & \frac{\lambda_{att}}{7} & \lambda_{task} \\ \lambda_{rec} & -(\lambda_{att} + \lambda_{task} + \lambda_{rec}) & 0 & 0 & 0 & \lambda_{att} & 0 & \lambda_{task} \\ \lambda_{rec} & 0 & -(\lambda_{att} + \lambda_{task} + \lambda_{rec}) & 0 & \lambda_{att} & 0 & 0 & \lambda_{task} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -(\lambda_{att} + \lambda_{task}) & 0 & \lambda_{rec} & \lambda_{task} \\ 0 & 0 & 0 & 0 & 0 & -(\lambda_{att} + \lambda_{task}) & \lambda_{rec} & \lambda_{task} \\ 0 & 0 & 0 & \frac{\lambda_{att}}{3} & \frac{\lambda_{att}}{3} & \frac{\lambda_{att}}{3} & -(\lambda_{att} + \lambda_{task}) & \lambda_{task} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

假设 GSPN 模型各个变迁实施速率相同且均为 1, 即 $\lambda_{att} = \lambda_{rec} = \lambda_{task} = 1$, 由式 (8) 和式 (9) 得模型各个状态稳态概率数值解为:

$$\begin{cases} P(M_0) = 0 \\ P(M_1) = 0 \\ P(M_2) = 0 \\ P(M_3) = 0.185 \\ P(M_4) = 0 \\ P(M_5) = 0 \\ P(M_6) = 0 \\ P(M_7) = 0.815 \end{cases} \quad (10)$$

本文基于蒙特卡洛仿真对系统木马攻击行为进行仿真, 设置仿真次数 10 000 次, 稳态概率结果对比如图 6 所示。

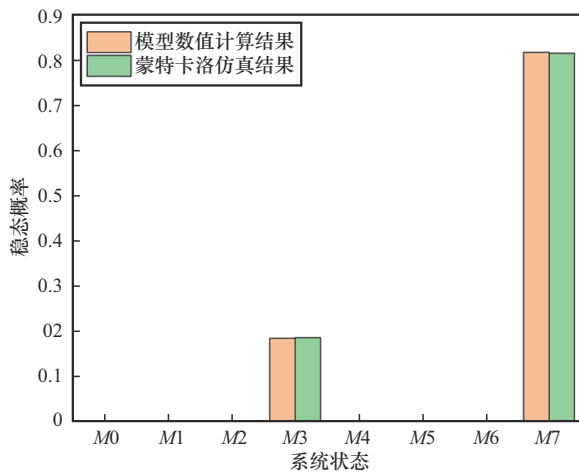


图6 结果对比

Figure 6 Comparison of results

由对比结果可知, 本文的模型数值计算结果和蒙特卡洛仿真结果误差仅为 0.001 83, 验证了

本文攻击模型和计算结果的准确性。

4 案例分析

4.1 实验设置

(1) 系统组成及功能

本文将系统执行的应用 DAG 图中的节点, 按其父子节点数量 (入度和出度) 分为 3 类: 父节点和子节点数量均小于等于 1; 父节点数大于 1, 称为“汇聚”节点; 子节点数大于 1, 称为“分散”节点 (一个节点既可以是“汇聚”节点也可以是“分散”节点)。为了降低实验复杂度并使结果具有一般性, 本节实验中假设案例应用 DAG 如图 7 所示, 在图 7 中, 同时具有 3 种类型的节点。

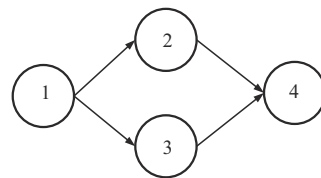


图7 案例应用 DAG

Figure 7 DAG of case application

(2) 安全评价指标

系统在运行时, 不同类型 HT 攻击对系统的安全性影响不同: 当被泄露信息 HT 攻击时, 系统可以正常运行, 此时系统的机密性受到破坏; 当被篡改输出木马攻击时, 系统虽然可以运行, 但是系统功能的完整性受到破坏; 当被拒绝服务木马攻击时, 系统中存在基本单元永久失效而导致可用性受到破坏。在本节后续实验中, 定义如下 3 个系统的安全性指标。

系统安全概率 (system security probability, SSP): 系统中执行任务的基本单元中 HT 处于休眠状态, 系统安全可用概率。

系统降级概率 (system degradation probability, SDP): 系统中基本单元被泄露信息或篡改输出木马攻击而导致系统机密性、完整性被破坏但仍然可用的概率。

系统失效概率 (system failure probability, SFP): 系统中执行任务的基本单元被拒绝服务木马攻击导致基本单元永久失效的概率。

(3) 对比实验设置

传统的基于检测的硬件木马安全评估方法从硬件的固有脆弱性即内因出发, 依据检测结果对硬件安全性进行评估, 安全性的评估只考虑了 HT 是否存在的静态问题, 缺乏对系统动态运行时 HT 是否触发从而对系统造成损害的考量。而基于本文 GSCPN 模型的安全性评估方法综合考虑内因和是否满足 HT 触发条件这一外因, 依据系统在运行时是否能够可靠安全地提供服务对安全性进行评估, 如系统中即使存在木马, 但未被触发, 则认为系统是安全的, 系统的安全性会随着系统的运行不断变化。两种方法的安全性分析的关注点和分析依据不同, 因此不便对其进行横向对比分析, 又因为本文是第一个进行 HT 攻击建模的研究, 缺乏实验对比, 所以本文只是在前一小节通过将数值解与仿真解进行比对说明了本文建模的准确性。在本节的后续实验中, 不将本文的方法和其他方法进行对比, 而是对比分析了不同攻击场景和系统任务场景下的系统安全性。

实验中参数设置及其含义如表4所示。

4.2 实验结果分析

本文通过更改参数值设置实验, 对比不同基本单元在不同木马攻击强度、不同木马恢复 (持续) 时间和不同任务执行时长对系统各种安全指标的影响, 并给出原因分析。

4.2.1 不同攻击强度对比

本小节通过改变参数 λ_{att}^i 进行仿真分析, 模拟不同 HT 攻击强度对系统安全性的影响, 如图8所示。当系统中任意一个基本单元的木马攻击强度增加时, SSP 降低, SFP 增加, SDP 在运行初始阶段也呈增加趋势。且基本单元1中木马攻击强度的增加对 SSP、SFP 和 SDP 影响幅度最大, 基本单元4中木马攻击强度的增加对 SSP、SFP 和 SDP 影响幅度最小。这是由于当基本单元中木马攻击强度增加时, 一方面增加了基本单元中一般木马激活概率和次数, 另一方面, 增加分布式协同木马触发逻辑条件满足的概率, 从而增加了其子单元中分布式木马负载逻辑激活概率, 因此 SSP 降低而 SDP 和 SFP 增加。又因为实验中基本单元1有两个子单元, 基本单元2和3只有一个子单元, 而基本单元4无子单元, 因此基本单元1中木马攻击强度的增加对 SSP、SDP 和 SFP 影响幅度最大, 基本单元2和3次之, 基本单元4中木马攻击强度的增加对 SSP、SDP 和 SFP 影响幅度最小。

4.2.2 不同恢复时间对比

本小节通过改变参数 λ_{rec}^i 进行仿真分析, 模拟不同 HT 恢复时间对系统安全性的影响, 如图9所示。当系统中任意一个基本单元的木马恢复时间增加时, 在运行初始阶段, 系统处于木马攻击状态时间增加, 所以 SSP 降低, SFP 和 SDP 增加。当处于稳定状态时, 相比基线实验, SSP 增加, SFP 降低, 这是由于木马恢复时间的增加降低了基本单元后续触发拒绝服务木马的概率, 提高了稳定状态的 SSP。但基本单元中木马恢复时间的增加对 SSP、SFP 和 SDP 的影响较为微弱, 在稳定状态下, SSP 最多增加约 0.002, SFP 最多降低约 0.002, 说明在基本单元平均硬件木马攻击时间为 10 h, 平均任务执行时间为 1 h 的场景下, 木马被激活后自动恢复至休眠状态的时间对系统的安全性影响不明显, 这是由于在该场景下, 基本单

表4 实验参数设置及含义

Table 4 Experimental parameter settings and meanings

参数	值	含义
λ_{att}^i	0.1、10	基本单元 i 中木马攻击的平均时间为 10 h 和 6 min, 模拟弱攻击场景和强攻击场景
λ_{rec}^i	0.1、10	基本单元 i 中木马处于被激活状态的平均时间为 10 h 和 6 min, 模拟木马长恢复时间和短恢复时间
λ_{task}^i	1、10	基本单元 i 执行一个任务的平均时间为 1 h 和 6 min, 模拟长时间任务和短时间任务

元执行任务期间激活木马概率较低, 所以后续触发拒绝服务木马的概率提升不明显。

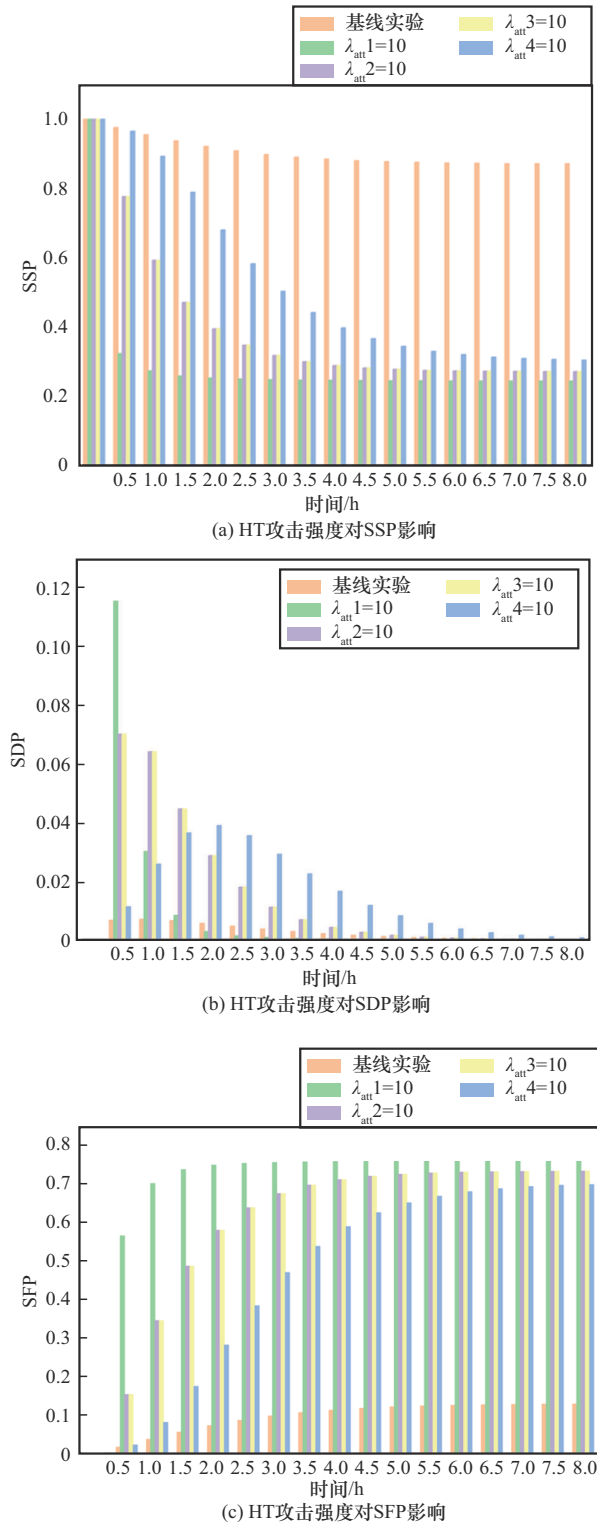


图8 不同基本单元不同木马攻击强度对安全性影响
Figure 8 Impact of different HT attack intensity on security for different basic units

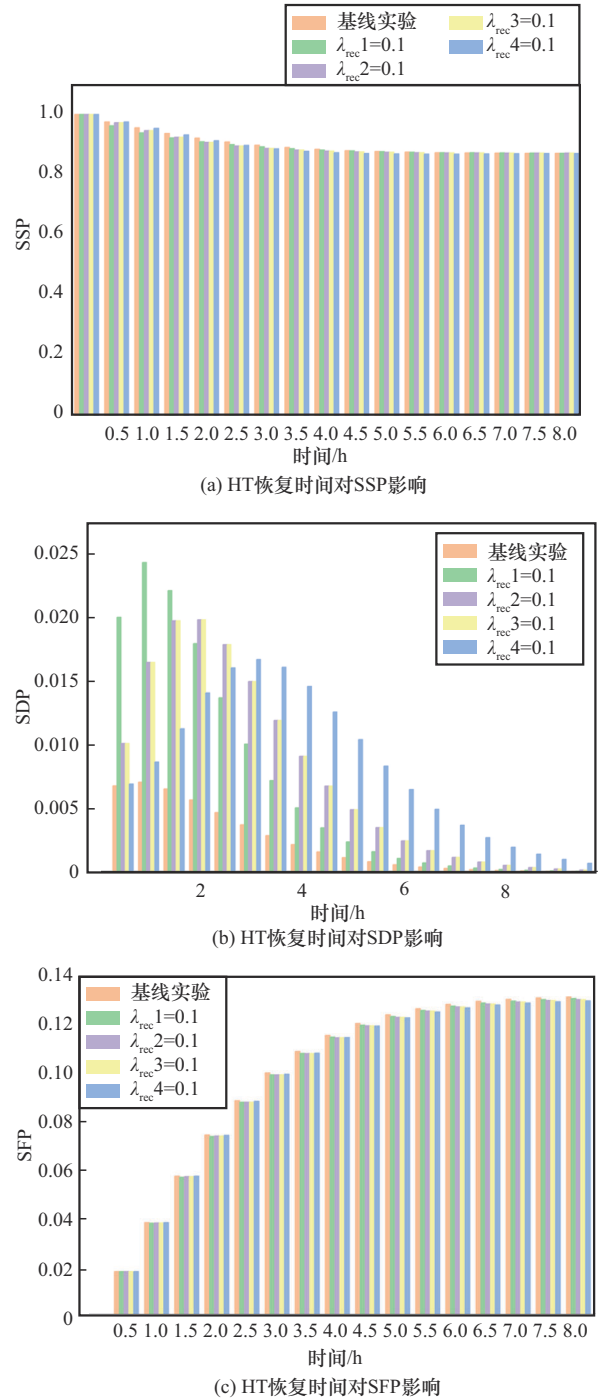


图9 不同基本单元不同木马恢复时间对安全性影响
Figure 9 Impact of different HT recovery time on security for different basic units

4.2.3 不同任务时长对比

本小节通过改变参数 $\lambda_{task}i$ 进行仿真分析, 模拟不同任务执行时间对系统安全性的影响, 如图10所示。系统中任意一个基本单元的任务执行时间减少, 则会降低运行期间木马触发条件满

足的概率和次数，减少基本单元一般木马激活和子单元分布式协同木马激活的概率，所以SSP增加，SFP降低。且基本单元1中任务运行时间减少对SSP、SFP和SDP影响幅度最大，基本单元2和基本单元3中任务执行时间减少对SSP、SFP和SDP影响幅度相同，基本单元4中任务执行时间减少对SSP、SFP和SDP影响幅度最小，原因和木马攻击强度分析部分相同。

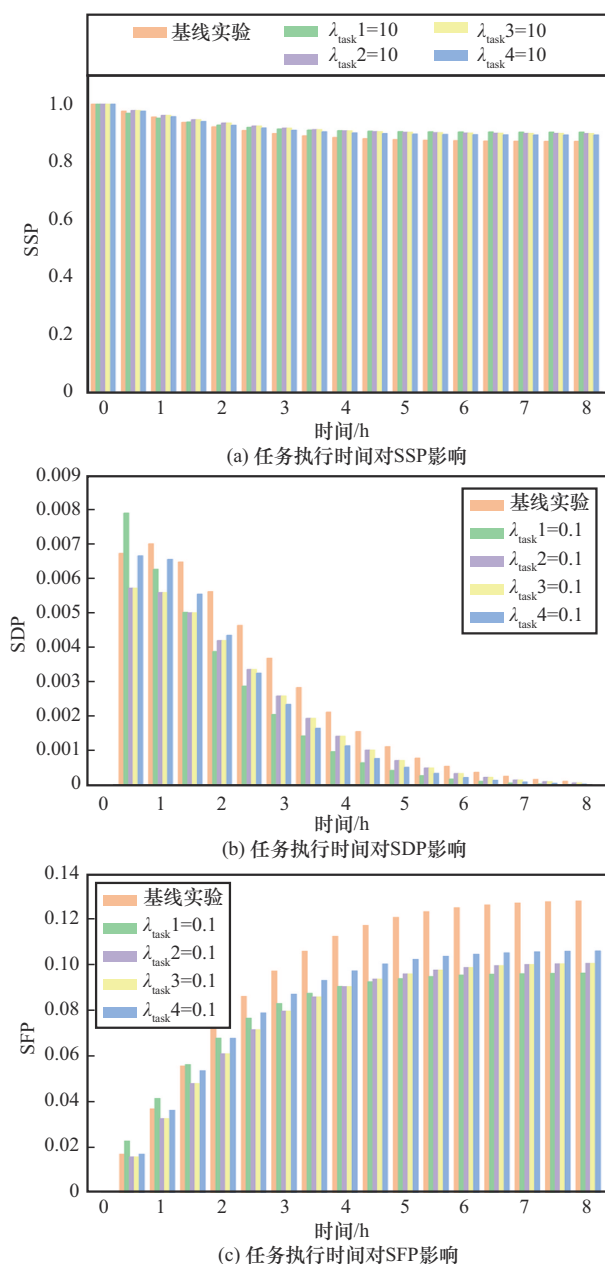


图10 不同基本单元不同任务执行时间对安全性影响

Figure 10 Impact of different task execution time on security for different basic units

5 结束语

本文提出了一种基于GSCPN的HT攻击建模方法，利用基本单元GSCPN木马攻击模型和一般性方法可以构建不同系统的攻击模型。本文的建模方法可以描述系统HT攻击的时序特性，描述系统状态变化与HT攻击之间的逻辑关系。通过求解与基本单元GSCPN模型同构的CTMC各个实存状态稳态分布的解析解和基于蒙特卡洛模拟方法的仿真解，二者的误差仅为0.001 83，验证了本文建模方法的准确性。通过简单的案例仿真实验表明，基本单元中HT攻击强度、HT恢复时间和任务执行时间对系统安全性有不同程度的影响。

本文所提的基于GSCPN的HT攻击模型可以为不同攻击场景下的系统安全性评估提供一个良好的解决方案。下一步将研究如何利用GSCPN刻画防御策略，构建攻防一体的系统安全评估模型，设计一种高安全增益的HT防御架构，并利用攻防一体化模型指导具体防御策略的制定。

参考文献：

- [1] AGRAWAL D, BAKTIR S, KARAKOYUNLU D, et al. Trojan detection using IC fingerprinting[C]//Proceedings of the 2007 IEEE Symposium on Security and Privacy, Piscataway: IEEE Press, 2007: 296-310.
- [2] CHEN Z, ZHOU P Q, HO T Y, et al. How secure is split manufacturing in preventing hardware Trojan? [C]//Proceedings of the 2016 IEEE Asian Hardware-Oriented Security and Trust (AsianHOST). Piscataway: IEEE Press, 2016: 1-6.
- [3] KHAN M S M, XI C J, KHAN A A, et al. Secure interposer-based heterogeneous integration[J]. IEEE Design & Test, 2022, 39(6): 156-164.
- [4] HU X, ZHAO Y, DENG L, et al. Practical attacks on deep neural networks by memory trojaning[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2021, 40(6): 1230-1243.
- [5] YASAEI R, FAEZI S, AL FARUQUE M A. Golden reference-free hardware trojan localization using graph convolutional network[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2022, 30(10): 1401-1411.
- [6] YANG S, CHAKRABORTY P, SLPSK P, et al. Trusted electronic systems with untrusted COTS[C]//Proceedings of the 2021 22nd International Symposium on Quality Electronic Design (ISQED). Piscataway: IEEE Press, 2021: 198-203.
- [7] PATNAIK S, ASHRAF M, SINANOGLU O, et al. A modern approach to IP protection and trojan prevention: split manufacturing for 3D ICs and obfuscation of vertical interconnects[J]. IEEE Transactions on Emerging Topics in Computing, 2021, 9(4): 1815-1834.
- [8] SAFARI Y, AGHANOURY P, IYER S S, et al. Hybrid obfuscation of chiplet-based systems[C]//Proceedings of the 2023 60th ACM/

- IEEE Design Automation Conference (DAC). Piscataway: IEEE Press, 2023: 1-6.
- [9] GUIMARÃES L A, BASTOS R P, FESQUET L. Detection of layout-level trojans by monitoring substrate with preexisting built-in sensors[C]//Proceedings of the 2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI). Piscataway: IEEE Press, 2017: 290-295.
- [10] DONG Z J, CHEN L, LI Y. A multi-level architecture for hardware Trojan and vulnerability runtime detection and response towards cryptographic IP[J]. IEICE Electronics Express, 2022, 19(11): 20220167.
- [11] CASSANO L, IAMUNDO M, LOPEZ T A, et al. DETON: DEfeating hardware Trojan horses in microprocessors through software obfuscation[J]. Journal of Systems Architecture, 2022, 129: 102592.
- [12] CHAKRABORTY R S, NARASIMHAN S, BHUNIA S. Hardware trojan: threats and emerging solutions[C]//Proceedings of the 2009 IEEE International High Level Design Validation and Test Workshop. Piscataway: IEEE Press, 2009: 166-171.
- [13] RAJENDRAN J J, SINANOGLU O, KARRI R. Building trustworthy systems using untrusted components: a high-level synthesis approach[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2016, 24(9): 2946-2959.
- [14] SUN Y D, JIANG G Y, LAM S K, et al. Designing energy-efficient MPSoC with untrustworthy 3PIP cores[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 31(1): 51-63.
- [15] RAO V V, SASAN A, SAVIDIS I. Analysis of the security vulnerabilities of 2.5-D and 3-D integrated circuits[C]//Proceedings of the 2022 23rd International Symposium on Quality Electronic Design (ISQED). Piscataway: IEEE Press, 2022: 1-7.
- [16] SLPSK P, RAY S, BHUNIA S. Treehouse: a secure asset management infrastructure for protecting 3DIC designs[J]. IEEE Transactions on Computers, 2023, 72(8): 2306-2320.
- [17] TEN C W, MANIMARAN G, LIU C C. Cybersecurity for critical infrastructures: attack and defense modeling[J]. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2010, 40(4): 853-865.
- [18] YU L, CHEN K L, CHANG Y, et al. A new correlation model of IoT attack based on attack tree[C]//Proceedings of the 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech). Piscataway: IEEE Press, 2021: 930-935.
- [19] 王赛娥, 刘彩霞, 刘树新, 等. 一种基于攻击树的4G网络安全风险评估方法[J]. 计算机工程, 2021, 47(3): 139-146, 154.
WANG S E, LIU C X, LIU S X, et al. A method of 4G network security risk assessment based on attack tree[J]. Computer Engineering, 2021, 47(3): 139-146, 154.
- [20] PHILLIPS C, SWILER L P. A graph-based system for network-vulnerability analysis[C]//Proceedings of the 1998 Workshop on New Security Paradigms. New York: ACM, 1998: 71-79.
- [21] 胡浩, 刘玉岭, 张玉臣, 等. 基于攻击图的网络安全度量研究综述[J]. 网络与信息安全学报, 2018, 4(9): 1-16.
HU H, LIU Y L, ZHANG Y C, et al. Survey of attack graph based network security metric[J]. Chinese Journal of Network and Information Security, 2018, 4(9): 1-16.
- [22] 原菊梅, 侯朝桢, 高琳, 等. 基于有色Petri网的复杂武器系统任务可靠性建模及估计[J]. 系统仿真学报, 2007, 19(11): 2563-2566.
YUAN J M, HOU C Z, GAO L, et al. Based on color petri-net modeling and estimate for task reliability of complex weapon system[J]. Journal of System Simulation, 2007, 19(11): 2563-2566.
- [23] ALMUTAIRI L M, SHETTY S. Generalized stochastic Petri Net

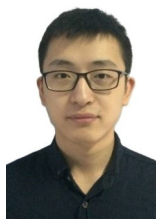
model based security risk assessment of software defined networks [C]//Proceedings of the MILCOM 2017 IEEE Military Communications Conference (MILCOM). Piscataway: IEEE Press, 2017: 545-550.

- [24] TANG S Y, HU X Y, ZHAO L. Modeling and security analysis of IEEE 802.1AS using hierarchical colored petri nets[C]//Proceedings of the GLOBECOM 2020 - 2020 IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-6.
- [25] BALBO G, CHIOLA G, FRANCESCHINIS G, et al. Generalized stochastic Petri nets for the performance evaluation of FMS[C]//Proceedings of 1987 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 1987: 1013-1018.

[作者简介]



梅波 (1998-), 男, 安徽宣城人, 信息工程大学博士生, 主要研究方向为内生安全、硬件木马防御。



郭威 (1990-), 男, 北京人, 博士, 信息工程大学副研究员, 主要研究方向为计算机体系结构、人工智能与大数据、内生安全。



曹志鹏 (1996-), 男, 陕西西安人, 信息工程大学博士生, 主要研究方向为晶上系统互连网络、计算机体系结构设计。



李沛杰 (1999-), 男, 山西襄汾人, 信息工程大学博士生, 助理研究员, 主要研究方向为硬件安全、软件定义晶上系统、软件定义互连。



张文博 (1999-), 男, 河南汝州人, 信息工程大学博士生, 主要研究方向为硬件安全、硬件木马。