行为经济学视域下的侧信道安全攻防策略博弈模型研究

蔡爵嵩 严迎建* 王晋东
(信息工程大学 郑州 450001)

摘 要：为解决侧信道安全领域暂无系统化攻防策略选择模型的现状，通过引入行为经济学和层次分析法，构建了一种侧信道攻防策略结构化成本量化模型。首先，从博弈论基本要素出发，将侧信道分析者与防护设计者映射为博弈双方，建立策略空间与收益函数。随后，提出基于层次分析法的成本量化框架，通过构建多维度判断矩阵，系统量化攻防策略的客观成本，克服传统主观赋值法的随意性与不可复现性。在此基础上，创新性地引入前景理论和风险规避理论，建立不完全信息静态博弈模型，并采用贝叶斯-纳什均衡求解最优策略。进一步结合安全工程ALARP原则，提出安全等级依赖的动态决策模型。结果表明，该模型能够为不同安全等级的场景提供基于风险偏好和成本效益原则量化的策略选择指导，明确识别出资产价值与防护强度的非线性临界关系，有效实现防护强度与资产价值的自适应匹配。通过构建“心理-经济-工程”三重维度的决策框架，不仅为侧信道攻防策略选择提供了首个系统化、可复现的理论模型，也为密码模块的“适度安全”设计提供了跨学科的量化决策依据。

关键词：侧信道安全；攻防策略；行为经济学；层次分析法；不完全信息静态博弈
中图分类号：TP309.7 **文献标识码**：A **文章编号**：1009-5896(2026)00-0001-12
DOI: 10.11999/JEIT260121 **CSTR**: 32379.14.JEIT260121

1 引言

侧信道分析与防护作为密码领域工程实践环节中的关键分析方法，相较于传统的密码理论分析，更具实践威胁性。通过对密码模块运行过程中产生的功耗[1]、电磁辐射[2]、时间[3]等物理量的变化分析，攻击者能够推断密码模块中的敏感信息，从而对密码系统构成实质性威胁。2024年，关于密码模块非入侵式缓解测试的相关国际标准ISO/IEC 17825-2024[4]和行业标准YD/T 6305-2024[5]的更新，进一步凸显了侧信道分析研究对密码芯片安全的重要性。然而，当前侧信道安全研究多集中于攻击方法与防护技术的创新，对于“攻防成本”这一影响策略选择的核心要素，尚未形成系统化、可量化的评估体系，制约了“适度安全”理念在工程实践中的落地。将侧信道分析与防护作为一个系统进行研究还未见公开文献，然而本文认为“攻防一体化”思想也应该引入侧信道分析与防护领域。侧信道分析与防护技术的应用落地，不能仅停留在技术革新上，还应该关注技术转化效益，才能进一步提高侧信道分析与防护的实用性。

针对侧信道安全与博弈论的融合研究，现有工作主要如下：姚等[6]在2012年提出了基于互信息博弈的侧信道安全风险评估模型，开创了博弈论在侧信道领域应用的先河。Wang等[7]对基于博弈论的云存储侧信道分析进行了研究，给出了云服务提供商在约束条件下的最优决策方法。Geng等[8]使用博弈论方法对随机动态电压缩放中如何选取电压阈值做了相应研究。以上文献都将博弈论作为策略选择的工具，但主要是针对侧信道分析与防护中的某个技术参数进行了优化，尚未涉及攻防策略的系统化成本量化问题，尤其是缺乏对“成本”本身进行结构化建模的方法论探讨。

虽然博弈论在侧信道领域中的应用研究极少，但博弈论理论早已深入到密码与安全领域的研究中了。2004年，Halpern等人[9]首次考虑到密码协议参与者的自利行为，将博弈论引入到了密码协议中，提出了理性密码协议的概念。2017年，彭长根等人[10]对博弈密码学进行了介绍，从博弈均衡的角度探讨了密码协议的公平性，并阐述了经济学中的机制设计及其在博弈密码协议设计中的应用及前景。

文献[11]指出，网络攻防与博弈论在目标对立、策略依存和关系非合作等特征方面具有相似性。正因如此，博弈论与网络攻防的交叉应用研究成果层出不穷。2025年，Sun等人[12]提出了网络攻防博弈模型，并借鉴传染病动力学定义网络态势，通过模拟不同策略组合，分析了网络零日病毒的宏观演化趋势。Bai等人[13]提出了基于马尔可夫信号博弈的计算机网络攻防均衡策略，有效增强对网络攻击的防御能力。Zheng等人[14]针对传统网络攻击预测器无法预测攻击时间的现状，提出了一种基于FlipIt的博弈论方法，不仅能够预测攻击时机，还能考虑攻防对抗的动态特性，自适应地提出防御策略。Zhu等人[15]在复杂的网络环境中，引入了传染

收稿日期：2026-xx-xx；改回日期：2026-06-30；网络出版：2026-07-13
*通信作者： 严迎建 yanyingjian@163.com

病模型，并结合FlipIt博弈，设计了一个复杂的网络攻击和防御时间博弈模型，通过动态调整时间策略，有效提升网络防御效能。当然，对于博弈论应用的研究，不仅仅局限于安全领域，只要涉及到策略选择，博弈论理论都能为研究人员提供一个有效的解决途径[16–20]。

上述研究虽然在策略优化方面取得了重要进展，但普遍建立在“完全理性”的经济人假设之上，即假设决策者总是最大化期望目标，且无偏地评估所有策略的成本与收益。然而，在实际安全决策中，决策者往往表现出有限理性，决策者的风险态度和损失厌恶心理会显著影响其对不同策略的偏好和选择。因此，单纯依赖客观成本评估无法完整刻画真实的策略选择过程，有必要引入行为经济学的理论工具，对决策者的心理因素进行建模，使博弈模型从“规范性分析”走向“描述性与规范性相结合”的现实决策分析。

2 博弈论在侧信道安全中的应用

2.1 侧信道安全在博弈论中的要素映射

按照《博弈论与经济行为》[21]提出的理论框架，由于侧信道安全中的攻防双方并没有一个具有约束力的协议，所以属于非合作博弈。如果按照局中人对博弈局势的知识即信息的了解程度，博弈问题又分为完全信息博弈和不完全信息博弈。本文认为，侧信道安全属于不完全信息静态博弈模型，即攻防双方在行动前无法确知对方将采取的具体策略类型，仅能依据先验信念进行决策。贝叶斯-纳什均衡(Bayes-Nash Equilibrium, BNE)是该场景下的求解方法，它描述了在给定先验信念下，每位局中人选择其最优策略的稳定状态。

具体地，将侧信道分析与防护问题采用博弈论的基本要素进行阐述：

(1)局中人：主要分为“攻击方”和“防御方”，其中“攻击方”为掌握侧信道分析方法的一方，“防御方”为掌握侧信道防护方法的一方。

(2)可行方案集(策略集)：假定局中人都是理性的，选择攻防策略时会同时考虑攻防结果和攻防成本(代价)两个方面的因素，所以在考虑进行攻防策略分类及其成本量化时也应该考虑这两个方面的因素。一般来说，攻防成本(代价)越高，攻防结果应该越好，这是攻防策略分类及其成本量化的基本原则。

(3)决策先后顺序：在侧信道分析与防护中，攻防双方决策有先后顺序，但后决策者(攻击方)对于先决策者(防御方)的策略只能通过经验或者尝试攻击方法才能逆向推导出来，所以其本质是静态博弈。

(4)支付(收益)函数：侧信道攻防是围绕目标密码模块的信息资产(密钥)而展开的，而攻防的效果是以攻防点信息资产安全风险的增减趋势予以判定的，并且取决于攻防双方的策略依存性。

(5)信息：在侧信道分析与防护中，信息是攻防双方有关博弈局势的知识，主要包括三个方面：一是目标的基本结构与功能、存在的脆弱性与威胁、已采取的安全防护措施等；二是攻防双方的可行方案集(策略集)，包括攻击策略集及相关攻击方法、防御策略集及相关的防御措施等；三是攻防双方支付(收益)函数的具体结构与数学表达式。

需要指出的是，本文将侧信道攻防建模为静态博弈，是基于当前侧信道安全评估的典型场景——即攻击者在有限次尝试内对固定防护配置的密码模块实施攻击，防御者在评估周期内策略不变。这种静态假设在安全性测评、产品认证等场景下具有合理性。然而，在更复杂的实际对抗中，攻击者可通过多次尝试逐步调整攻击参数(如增加能量迹数量、优化预处理方式)，防御者亦可依据安全日志或测评反馈迭代升级防护措施，这更接近动态博弈或重复博弈的框架。本文以静态模型为基础进行分析，为后续动态化扩展为动态博弈或重复博弈框架奠定理论与方法论基础。

2.2 侧信道安全的博弈模型结构

在侧信道分析与防护中，攻防博弈模型是攻防双方根据所掌握的有关博弈局的知识(信息)而展开博弈过程的描述，博弈对象是目标密码模块并主要针对其脆弱性展开攻防博弈。首先，必须明确的是，脆弱性是客观存在的，在攻防演化过程中目标的脆弱性会不断发生变化，攻防双方据此选择各自的攻防策略，努力使各自的支付(收益)达到最大。同时，攻防双方在实施各自的攻防策略时，必须考虑对方策略的影响，所以攻防博弈模型必须同时包含攻击策略和防御策略，而不能只从单方面进行考虑，即攻击方受到防御方的影响，防御方受到攻击方的影响，这是博弈模型的基本假定。

对于攻击方而言，通过提高攻击能力，运用各种侧信道分析技术，发现、利用目标的脆弱性知识，增加其安全风险，增大攻击成功的可能性，但同时受防御方策略的影响，所以攻击效果具有不确定性。对于防御方而言，通过提高防御水平，运用各种侧信道防护技术，发现、弥补目标的脆弱性以降低其安全风险，增大防御成功的可能性，但同时受攻击方策略的影响，所以防御效果也具有不确定性。

根据侧信道分析与防护双方的博弈过程，本文针对目标模块的侧信道分析与防护构建侧信道攻防博弈模型结构，如图1所示。

在构建博弈模型时，策略成本是决定收益函数与均衡结果的关键输入。传统研究中将成本简化为单一经济指标，难以反映策略的技术复杂度与资源依赖特性。为此，本文在第3章将攻击与防护策略的成本分解为多维度准则，并通过判断矩阵计算权重，从而为博弈模型提供可解释、可比较、可调整的成本参数。

3　多层次融合的侧信道攻防策略结构化成本量化框架

传统研究中攻防成本常被简化为单一数值或依赖专家直接赋值，缺乏结构性与可解释性。为此，本文首次将层次分析法(Analytic Hierarchy Process, AHP)[22][23]系统引入侧信道安全领域，构建从“准则分解”到“权重计算”再到“行为调整”的多层次成本量化框架，使成本量化从单一工程维度扩展至“心理-经济-工程”三重维度。

具体地，本文首先通过构建层次化的成本维度，并结合1-9标度法进行两两比较，将决策者的经验判断转化为可计算的权重，从而对复杂决策问题进行层次化数值分析，这对于还没有统一“成本-收益”标准的侧信道安全领域是十分重要的。然后引入前景理论[24]和风险规避理论[25]来调整AHP量化成本，最终使用安全工程中的ALARP (As Low As Reasonably Practicable)原则进一步约束成本，使博弈模型能够更真实地反映实际决策中的心理因素和工程约束。

3.1　侧信道分析策略及其成本构成

侧信道分析策略根据是否需要对泄露信息进行先验建模，可分为非建模类分析与建模类分析。本文选取最具代表性和实际威胁的四种分析技术构成攻击方的策略集$S_A=\{a,b,c,d\}$：

(1)策略a为“不攻击”，作为基准策略，其成本与成功率均为0。

(2)策略b为“DPA/CPA”，属于非建模类分析，通过统计分析大量能量迹与中间值似之间的相关性来恢复密钥，其成本构成包括：

$$C_b = C_{\text{time}} + C_{\text{data}} + C_{\text{computation}} \tag{1}$$

其中，C_{time}为数据采集时间成本，C_{data}为存储成本，$C_{\text{compotation}}$为计算成本。

(3)策略c为“模板攻击TA”，属于建模类分析，需要针对目标设备构建泄露模板，分析时通过模板匹配进行分类，其成本构成包括：

$$C_c = C_b + C_{\text{modeling}} + C_{\text{expertise}} \tag{2}$$

其中，C_{modeling}为建模成本，$C_{\text{expertise}}$为专家知识成本。

(4)策略d为“基于深度学习的侧信道分析DLSCA”，也属于建模类分析，利用深度学习模型自动提取特征并进行分类，其成本构成包括：

$$C_d = C_c + C_{\text{GPU}} + C_{\text{training}} + C_{\text{high-tech}} \tag{3}$$

其中，C_{GPU}为GPU算力成本，C_{training}为模型训练成本，$C_{\text{high-tech}}$为高技术门槛成本。

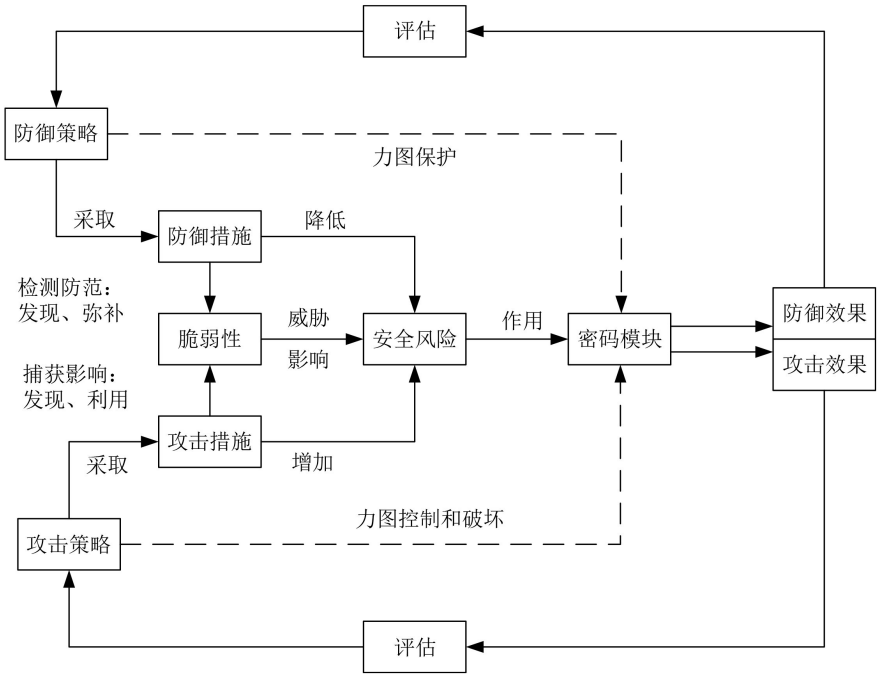


图 1 侧信道安全博弈模型结构

注意：策略成本呈现超线性增长特征，即$C_d - C_c > C_c - C_b > C_b - C_a$，这反映了高级分析技术面临显著的边际成本递增效应。

3.2 侧信道防护策略及其成本构成

根据GM/T 0083-2020[26]，本文将防御方策略集S_D={A,B,C,D}定义如下：

(1)策略A为“无防护”，作为基准策略，不采取任何防护措施。

(2)策略B为“时间隐藏技术”，通过在算法执行中引入随机时延、乱序操作等手段，使操作时间与处理数据无关，其成本构成为：

$$C_B = C_{\text{microarch}} + C_{\text{design}} + C_{\text{performance}} \tag{4}$$

其中，$C_{\text{microarch}}$为微架构改动成本，C_{design}为设计复杂度成本，$C_{\text{performance}}$为性能损失成本。

(3)策略B为“振幅维度隐藏技术”，通过双轨预充电逻辑、增加噪声、低功耗设计等方法，使能量消耗振幅与处理数据无关，其成本构成为：

$$C_C = C_B + C_{\text{circuit}} + C_{\text{area}} + C_{\text{power}} \tag{5}$$

其中，C_{circuit}为电路级设计成本，C_{area}为面积开销成本，C_{power}为功耗增加成本。

(4)策略C为“掩码/盲化技术”，通过在算法运算过程中引入随机数(掩码)或随机化中间值(盲化)，从算法设计层面根本上隐藏真实数据，其成本构成为：

$$C_D = C_C + C_{\text{algorithm}} + C_{\text{verification}} + C_{\text{crypto-expertise}} \tag{6}$$

其中，$C_{\text{algorithm}}$为算法修改成本，$C_{\text{verification}}$为验证复杂度成本，$C_{\text{crypto-expertise}}$为密码学专业知识成本。

注意：本文假设防护策略成本呈现出几何级数增长特征。设基础成本为C_0，则有：

$$C_A = C_0,\ C_B = k_1 C_0,\ C_C = k_1 k_2 C_0,\ C_D = k_1 k_2 k_3 C_0 \tag{7}$$

其中，k_1、k_2、k_3为增长系数。本文认为随着防护措施强度的增加，后续的防护措施成本应该在之前的防护措施成本的基础上进行增加。

3.3 基于AHP的攻防策略成本结构化量化

3.3.1 传统AHP成本量化

基于1-9标度法，构建攻防双方的准则层判断矩阵如下：

(1)攻击方准则层判断矩阵M_A^{crit}及权重W_A^{crit}：

$$M_A^{\text{crit}} = \begin{bmatrix} 1 & 1/3 & 1/5 & 1/2 \\ 3 & 1 & 1/3 & 2 \\ 5 & 3 & 1 & 3 \\ 2 & 1/2 & 1/3 & 1 \end{bmatrix} \Rightarrow$$

$$W_A^{\text{crit}} = \begin{bmatrix} 0.085 & 0.240 & 0.523 & 0.152 \end{bmatrix}^T \tag{8}$$

(2)防御方准则层判断矩阵M_D^{crit}及权重W_D^{crit}：

$$M_D^{\text{crit}} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1/2 & 1 & 2 & 3 \\ 1/3 & 1/2 & 1 & 2 \\ 1/4 & 1/3 & 1/2 & 1 \end{bmatrix} \Rightarrow$$

$$W_D^{\text{crit}} = \begin{bmatrix} 0.467 & 0.277 & 0.160 & 0.095 \end{bmatrix}^T \tag{9}$$

(3)通过层次总排序，得到攻防双方的综合权重向量：

$$W_A^{\text{com}} = \begin{bmatrix} 0.047 & 0.122 & 0.265 & 0.566 \end{bmatrix}^T \tag{10}$$

$$W_D^{\text{com}} = \begin{bmatrix} 0.050 & 0.121 & 0.264 & 0.565 \end{bmatrix}^T \tag{11}$$

3.3.2 行为经济学调整与ALARP原则约束

本文认为单纯使用AHP来量化成本，不能考虑攻防过程中决策者的心理因素变化，故引入主观价值调整函数：

$$c_{\text{subj}}(s, SL) = \rho(SL) \times [c_{\text{AHP}}(s) + \beta \times \Delta c_{\text{behavioral}}(s)] \tag{12}$$

其中，$\rho(SL)$为安全等级调整系数，反映不同场景下的价值感知差异；β为行为因素权重，$\beta \in [0,1]$；$\Delta c_{\text{behavioral}}(s)$为行为成本调整项，定义如下：

$$\Delta c_{\text{behavioral}}(s) = \lambda(SL) \times I_{\text{loss}}(s) + \alpha(SL) \times \sigma_s^2 \tag{13}$$

其中，$I_{\text{loss}}(s)$为策略s可能带来的损失感知，σ_s^2为策略s的收益方差(风险度量)。

因为侧信道分析与防护是一个工程实践性很强的研究内容，所以本文进一步引入了安全工程中的ALARP原则来约束成本：

$$\frac{c(s)}{V \times \Delta R(s)} \in [\theta_{\min}(SL), \theta_{\max}(SL)] \tag{14}$$

其中，$\Delta R(s)$为策略s带来的风险降低值，$\theta_{\min}(SL)$和$\theta_{\max}(SL)$为安全等级依赖的成本效益比边界。

3.3.3 多层次融合的成本计算公式

综合考虑上述因素，得到最终的成本计算公式：

$$c_{\text{final}}(s, V, SL) = \begin{cases} 0 & \text{if} \dfrac{c_{\text{subj}}(s, SL)}{V \times \Delta R(s)} < \theta_{\min}(SL) \\ c_{\text{subj}}(s, SL) & \text{if} \dfrac{c_{\text{subj}}(s, SL)}{V \times \Delta R(s)} \in [\theta_{\min}(SL), \theta_{\max}(SL)] \\ \infty & \text{if} \dfrac{c_{\text{subj}}(s, SL)}{V \times \Delta R(s)} > \theta_{\max}(SL) \end{cases} \tag{15}$$

4 基于不完全信息静态博弈的侧信道安全攻防策略模型

4.1 模型形式化定义

4.1.1 基本要素

本文构建的行为增强贝叶斯博弈模型可形式化定义为六元组：

$$\mathcal{G} = \langle \mathcal{N}, \mathcal{S}, \Theta, \mathcal{P}, \mathcal{U}, \mathcal{B} \rangle \tag{16}$$

其中：(1)$\mathcal{N} = \{A, D\}$为局中人集合，A表示攻击方，D表示防御方。(2)$\mathcal{S} = S_A \times S_D$为策略空间，$S_A$=$\{a,b,c,d\}$为攻击策略集，$S_D$=$\{A,B,C,D\}$为防御策略集。(3)$\Theta = \Theta_A \times \Theta_D$为类型空间，$\Theta_i$表示局中人$i$的私有类型，包含其风险偏好、损失厌恶等行为特征。(4)$\mathcal{P} = \{p_A, p_D\}$为成功率函数，$p_A(s_a,s_d)$表示攻击方采用策略$s_a$对抗防御策略$s_d$的成功率，$p_D$=1-$p_A$。(5)$\mathcal{U} = \{u_A, u_D\}$为行为调整的效用函数，具体定义见公式(24)。(6)$\mathcal{B} = \{\beta_A, \beta_D\}$为信念系统，$\beta_A \in \Delta(S_D)$为攻击方对防御方策略的先验信念，$\beta_D \in \Delta(S_A)$为防御方对攻击方策略的先验信念。其中，$\Delta(X)$表示$X$上的概率分布。

4.1.2 类型依赖的行为参数

局中人i的类型$\theta_i \in \Theta_i$决定了其行为参数：

$$\theta_i = (\alpha_i, \lambda_i, \rho_i, R_{0,i}) \tag{17}$$

其中：(1)α_i为风险规避系数，描述了决策者面对不确定性时的偏好特征：$\alpha_i > 0$表示风险规避，$\alpha_i = 0$表示风险中性，$\alpha_i < 0$表示风险偏好。(2)λ_i为损失厌恶系数，描述了决策者对损失的敏感度高于对等额收益的敏感度，通常大于1。(3)ρ_i为安全等级调整系数。(4)$R_{0,i}$为前景理论中的“参考点”。这些参数与密码模块安全等级相关，并满足映射关系：$\theta_i = f_i(SL)$。

4.2 行为增强的效用函数

(1)首先定义基础效用函数(不考虑行为因素)：

$$u_A^0(s_a, s_d) = p_A(s_a, s_d) \cdot V - c_A(s_a) \tag{18}$$

$$u_D^0(s_d, s_a) = p_D(s_d, s_a) \cdot V - c_D(s_d) \tag{19}$$

其中，V为资产价值，$c_A(s_a)$和$c_D(s_d)$为第3章量化的策略成本。

(2)考虑决策者的风险偏好，采用指数效用函数进行风险调整：

$$u_A^{\text{risk}}(s_a, s_d) = \begin{cases} \dfrac{1 - e^{-\alpha_A \cdot u_A^0(s_a, s_d)}}{\alpha_A} & \text{if } \alpha_A \neq 0 \\ u_A^0(s_a, s_d) & \text{if } \alpha_A = 0 \end{cases} \tag{20}$$

$$u_D^{\text{risk}}(s_d, s_a) = \begin{cases} \dfrac{1 - e^{-\alpha_D \cdot u_D^0(s_d, s_a)}}{\alpha_D} & \text{if } \alpha_D \neq 0 \\ u_D^0(s_d, s_a) & \text{if } \alpha_D = 0 \end{cases} \tag{21}$$

(3)基于参考点R_0进行前景理论调整，考虑损失厌恶效应：

$$u_A^{\text{PT}}(s_a, s_d) = \begin{cases} u_A^{\text{risk}}(s_a, s_d) - R_{0,A} & \text{if } u_A^{\text{risk}} \geq R_{0,A} \\ \lambda_A \times (u_A^{\text{risk}}(s_a, s_d) - R_{0,A}) & \text{if } u_A^{\text{risk}} < R_{0,A} \end{cases} \tag{22}$$

$$u_D^{\text{PT}}(s_d, s_a) = \begin{cases} u_D^{\text{risk}}(s_d, s_a) - R_{0,D} & \text{if } u_D^{\text{risk}} \geq R_{0,D} \\ \lambda_D \times (u_D^{\text{risk}}(s_d, s_a) - R_{0,D}) & \text{if } u_D^{\text{risk}} < R_{0,D} \end{cases} \tag{23}$$

(4)综合考虑ALARP原则，定义最终效用函数：

$$u_{\text{i}}^{\text{final}}(s_i, s_{-i}) = \begin{cases} u_i^{\text{PT}}(s_i, s_{-i}) & \text{if } \dfrac{c_i(s_i)}{V \cdot \Delta \text{R}_i(\text{s}_i)} \in \Theta_i^{\text{ALARP}} \\ -\infty & \text{otherwise} \end{cases} \tag{24}$$

其中，$\Theta_i^{\text{ALARP}} = [\theta_i^{\min}, \theta_{\text{i}}^{\max}]$为ALARP可接受区间。公式(24)表明，效用函数不再是简单的线性函数，而是分段非线性函数，这更真实地反映了实际决策过程。

4.3 贝叶斯-纳什均衡求解

(1)在不完全信息下，局中人基于信念计算期望效用如下：

(a)防御方选择策略s_d的期望效用为：

$$EU_D(s_d; \sigma_A, \beta_D) = \sum_{s_a \in S_A} \beta_D(s_a) \cdot u_D^{\text{final}}(s_d, s_{\text{a}}) \tag{25}$$

(b)攻击方选择策略s_a的期望效用为：

$$EU_A(s_a; \sigma_D, \beta_A) = \sum_{s_d \in S_D} \beta_A(s_d) \cdot u_A^{\text{final}}(s_{\text{a}}, s_d) \tag{26}$$

(2)混合策略组合(σ_A^*, σ_D^*)构成贝叶斯-纳什均衡，如果对于所有$s_a \in S_A$和$s_d \in S_D$，满足：

$$\sigma_A^*(s_a) > 0 \Rightarrow EU_A(s_a; \sigma_D^*, \beta_A) = \max_{s_{a'} \in S_A} EU_A(s_{a'}; \sigma_D^*, \beta_A) \tag{27}$$

$$\sigma_D^*(s_d) > 0 \Rightarrow EU_D(s_d; \sigma_A^*, \beta_D) = \max_{s_{d'} \in S_D} EU_D(s_{d'}; \sigma_A^*, \beta_D) \tag{28}$$

(3)本文提出基于行为增强的贝叶斯-纳什均衡求解算法求解行为增强的BNE。

4.4 模型的理论性质

定理4.1(均衡存在性)：对于任意有限策略空间S_A和S_D，任意连续的行为调整效用函数$u_{\text{i}}^{\text{final}}$，以及任意信念分布$\{\beta_A, \beta_D\}$，行为增强的贝叶斯博弈$\mathcal{G}$至少存在一个混合策略贝叶斯-纳什均衡。

证明：根据Glicksberg不动点定理，由于策略空间是有限的离散空间，效用函数是连续的，且混合策略空间是紧凸集，因此存在不动点即纳什均衡。

定理4.2(资产价值单调性)：设$(\sigma_D^*(V), \sigma_A^*(V))$为资产价值$V$下的BNE，则对于防御方，存在阈值$V_{\text{th}}$使得：

$$\frac{\partial \mathbb{E}[\sigma_D^*(V)(s_d)]}{\partial V} > 0 \,\text{for} s_d \in \{B, C, D\} \tag{29}$$

表 1 行为增强的贝叶斯-纳什均衡求解算法

算法1 行为增强的贝叶斯-纳什均衡求解算法

Input：资产价值V，安全等级SL，成本矩阵$\boldsymbol{C}$，成功率矩阵P，信念分布$\{\beta_A,\beta_D\}$

Output：均衡策略(σ_A^*,σ_D^*)

1 确定行为参数：$\alpha_D(\mathrm{SL}),\alpha_A(\mathrm{SL}),\lambda_D(\mathrm{SL}),\lambda_A(\mathrm{SL}),R_{0,D},R_{0,A}$；

2 计算调整后成本：$\boldsymbol{C'}\leftarrow \mathrm{AdjustCost}(\boldsymbol{C},\mathrm{SL},V)$ //公式(15)

3 计算风险调整收益：$U^{\mathrm{risk}}\leftarrow \mathrm{RiskAdjust}(\boldsymbol{P},V,\boldsymbol{C'},\alpha_D,\alpha_A)$

4 应用前景理论调整：$U^{\mathrm{PT}}\leftarrow \mathrm{ProspectAdjust}(U^{\mathrm{risk}},\lambda_D,\lambda_A,R_{0,D},R_{0,A})$

5 应用ALARP约束：$S_{\mathrm{feasible}}\leftarrow \mathrm{ALARPFiliter}(U^{\mathrm{PT}},\boldsymbol{C'},V,\Theta^{\mathrm{ALARP}})$

6 构建优化问题：

7 目标函数：$\max_{\sigma_D,\sigma_A}\left(\sum_{s_d}\sigma_D(s_d)\cdot \mathrm{EU}_D(s_d)+\sum_{s_a}\sigma_A(s_a)\cdot \mathrm{EU}_A(s_a)\right)$

8 约束条件：$\sum\sigma_D=1,\sum\sigma_A=1,\sigma_i\geq 0,s_i\in S_{\mathrm{feasible}}$

9 使用SLSQP算法求解上述优化问题；

10 验证BNE条件：检查是否满足公式(27)和(28)；

11 If 验证失败 then

12 | 调整初始点，重新执行步骤7-9求解优化问题；

13 end

14 return (σ_A^*,σ_D^*)；

即随着资产价值增强，防御方采用防护策略的概率单调增强。

定理4.3(安全等级效应)：设$(\sigma_D^*(SL),\sigma_A^*(SL))$为安全等级SL下的BNE，则有：

$$\frac{\partial\alpha_D}{\partial SL}<0,\ \frac{\partial\lambda_D}{\partial SL}<0\Rightarrow\frac{\partial\mathbb{E}[\sigma_D^*(SL)(D)]}{\partial SL}>0 \tag{30}$$

即随着安全等级提升，防御方的风险规避和损失厌恶减弱，采用高级防护策略的概率增加。

5 实验与分析

为全面验证本文提出的行为增强博弈模型的有效性、鲁棒性与实用价值，本章设计并实施了系统性的实验评估。实验围绕三个核心目标展开：1.验证基于AHP与行为经济学的成本量化模型的合理性与一致性；2.分析模型在不同安全等级与资产价值下的贝叶斯-纳什均衡特征；3.通过敏感性分析评估模型对关键参数(如资产价值)的敏感性。所有实验均基于第3章的成本量化模型和第4章的博弈模型，并在以下环境进行：Intel i9-13900HX CPU，96GB RAM， Windows 11 64-bit， Python 3.10。BNE求解使用SciPy库的SLSQP算法，设置收敛容差ϵ=10^{-8}。

5.1 实验参数设定

实验参数综合参考国际标准ISO/IEC 17825-2024、行业标准GM/T 0083-2020[26]，并结合专家经验确定。基准参数设置如下：(1)资产价值V：基准值V=1.0，敏感性分析范围$V\in[0.1,5.0]$。(2)攻击成功率矩阵PA：基于技术成熟度和专家评估，设定如表2所示数值。

(3)先验信念β：基于标准推荐和行业实践，设定如公式(31)(防御方对攻击方的信念)和(32)(攻击方对防御方的信念)所示数值。

$$\beta_D=[0.2,0.5,0.2,0.1] \tag{31}$$

$$\beta_A=[0.3,0.4,0.2,0.1] \tag{32}$$

(4)行为参数：基于密码模块安全等级，设定如表3所示数值。

5.2 成本量化模型验证

图2对比了传统AHP模型与行为增强模型在不同安全等级下的成本曲线。

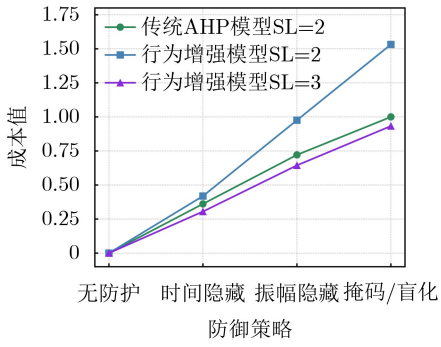
由图2可知，(1)在低安全等级(SL=1)的情况下，由于高风险规避(α大)和高损失厌恶(λ大)，决策者感知的成本被显著放大。在高安全等级(SL=3)，行为调整效应减弱，成本感知更接近客观AHP成本。(2)攻击方与防御方的成本曲线均随安全等级升高而升高，并且斜率也随之增大，说明高安全等级下，攻防双方的斗争更加激烈。

表 2 攻击成功率矩阵$P_A(s_a,s_d)$

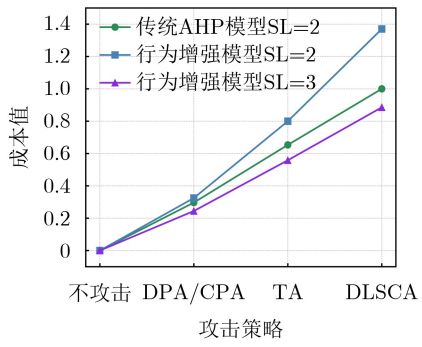
p_{A}\防御策略	A(无防护)	B(时间隐藏)	C(振幅隐藏)	D(掩码/盲化)
a(不攻击)	0.00	0.00	0.00	0.00
b(DPA/CPA)	0.60	0.40	0.10	0.20
c(TA)	0.80	0.10	0.25	0.30
d(DLSCA)	0.90	0.30	0.35	0.50

表 3　安全等级依赖的行为参数设定

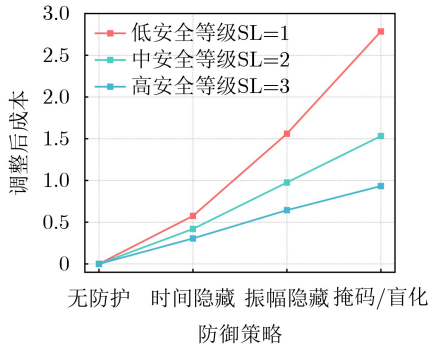
安全等级	风险规避系数α		损失厌恶系数λ		ALARP边界$[\theta_{\min},\theta_{\max}]$		ρ	参考点R_0
SL=1	1.50(攻)	1.20(防)	2.50(攻)	2.00(防)	[0.05, 0.10](攻)	[0.10, 0.20](防)	1.20	0
SL=2	0.80(攻)	0.60(防)	1.80(攻)	1.50(防)	[0.10, 0.30](攻)	[0.15, 0.40](攻)	1.00	0
SL=3	0.30(攻)	0.20(防)	1.20(攻)	1.00(防)	[0.30, 0.80](攻)	[0.25, 0.60](攻)	0.80	0



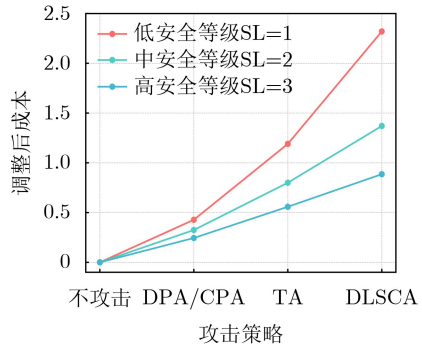
(a) 防御方在不同模型下的成本对比



(b) 攻击方在不同模型下的成本对比



(c) 防御方在不同安全等级下的成本对比



(d) 攻击方在不同安全等级下的成本对比

图 2 成本曲线对比

为验证基于AHP的成本结构化量化模型的有效性，本文对所有判断矩阵进行一致性检验，确保主观判断的逻辑合理性。

为进一步评估AHP权重在输入判断矩阵微小扰动下的可靠性，本文定义了“权重稳定性”指标，通过计算权重向量的变异系数来度量各维度权重分布的均匀性，计算公式如下：

$$\text{Stability} = 1 - CV = 1 - (\sigma(w)/\mu(w)) \tag{33}$$

其中，$\sigma(w)$和$\mu(w)$分别表示权重向量w的标准差和均值。当该指标值越接近1，表明各维度权重分布越均匀，对判断矩阵的敏感性越低，即稳定性越高；反之，若权重分布极不均匀，则表明成本结构高度依赖于某一特定准则，稳定性相对较低。

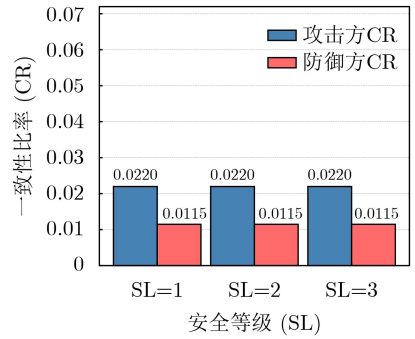
图3中的权重是由AHP判断矩阵计算得出的，体现了各成本维度的相对重要性，为后续成本综合计算提供了结构化基础。具体来说，图3中(a)展示了三种安全等级下攻击方与防御方准则层判断矩阵的一致性比率(CR)，可以看出，所有CR均小于0.1，表明本文假设的数值具有良好的一致性，成本量化的主观判断过程是可靠的。图3中(b)展示了权重稳定性分析结果，可以看出，虽然稳定性指标并未达到高稳定性水平，但从侧面印证了本文引入行为经济学调整的必要性：当客观成本评估存在一定不确定性时，决策者的心理因素将对最终决策产生更显著的影响。为进一步说明准则层权重的分布特征，本文以中安全等级(SL=2)为例进行分析。图3中(c)和(d)展示了攻击方与防御方在SL=2下的准则层权重分布，可见AHP成功将模糊的“成本”概念转化为可解释的权重向量，如攻击方重视“技术门槛”，而防御方侧重“研发设计”，符合领域直觉。

5.3　贝叶斯-纳什均衡分析

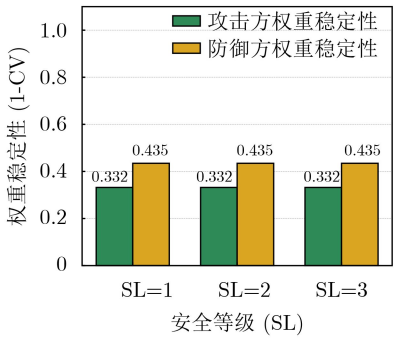
下面计算不同场景下的贝叶斯-纳什均衡，图4分别展示了三个不同安全等级下的均衡策略概率分布。

从不同安全等级下的BNE策略概率分布图可以看出：(1)对于低价值资产，在强风险规避心理

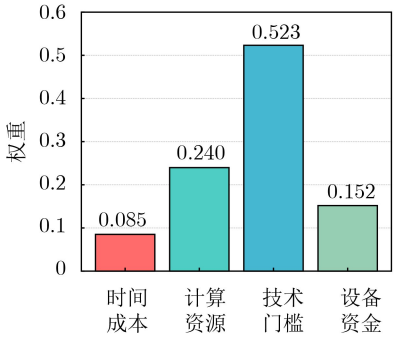
下，双方均倾向于“不作为”以规避任何潜在损失，符合ALARP原则的下界。(2)高价值资产驱动双方采用高技术、高成本的策略，追求高成功率或高强度防护。



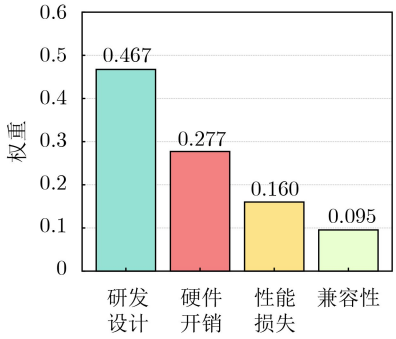
(a) 不同安全等级下AHP一致性检验结果



(b) 不同安全等级下权重稳定性结果

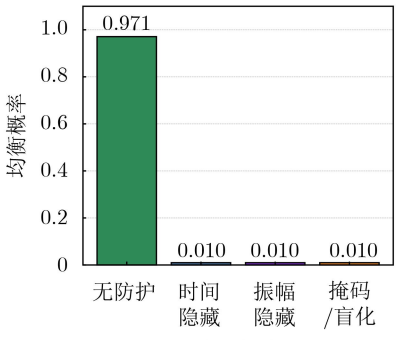


(c) 中安全等级下攻击方成本权重分布

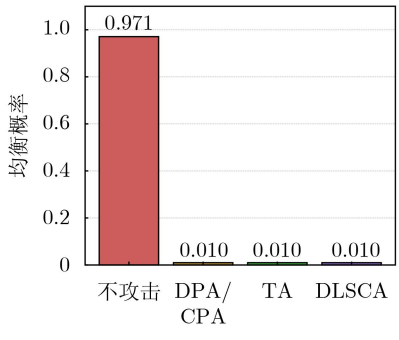


(d) 中安全等级下防御方成本权重分布

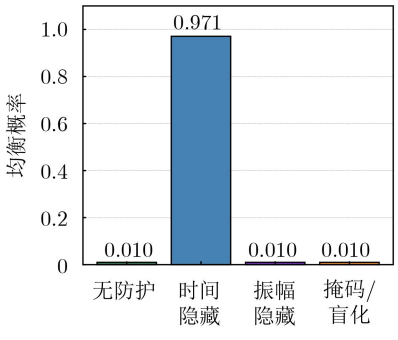
图 3 AHP成本量化一致性检验与权重分析



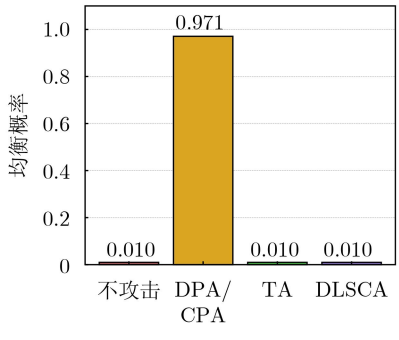
(a) 低安全等级下防御方BNE策略概率分布



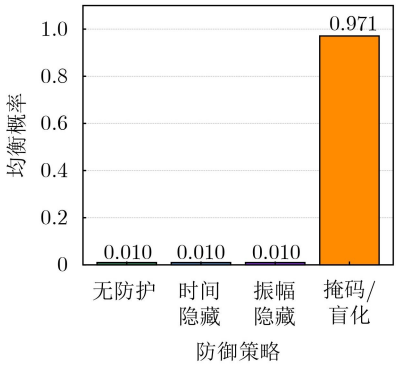
(b) 低安全等级下攻击方BNE策略概率分布



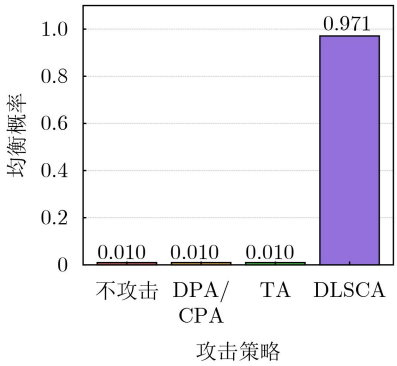
(c) 中安全等级下防御方BNE策略概率分布



(d) 中安全等级下攻击方BNE策略概率分布



(e) 高安全等级下防御方BNE策略概率分布

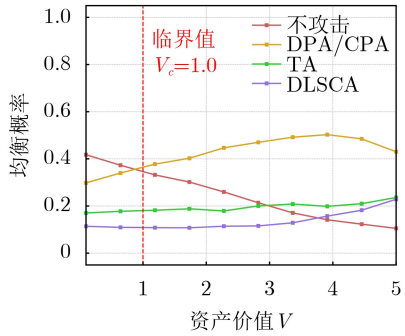


(f) 高安全等级下攻击方BNE策略概率分布

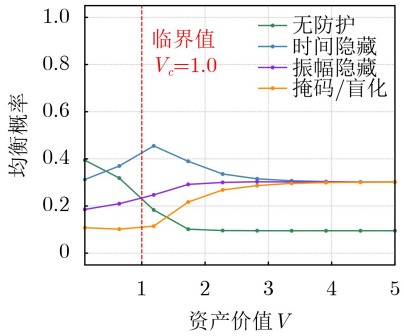
图 4 不同安全等级下攻防双方BNE策略概率分布

5.4 敏感性分析

为深入探究模型关键参数对均衡策略的影响，本节进行系统的敏感性分析。图5展示了在中安全等级(SL=2)下，攻防双方均衡策略概率随资产价值V变化的关系曲线。可以看出，在资产价值$V_{th} \approx 1.0$附近，攻防双方的策略分布均发生显著变化，这与定理4.2(资产价值单调性)的预测一致，即攻防策略概率会随资产价值增加而进行调整。



(a) 攻击方均衡策略概率随资产价值V的变化趋势



(b) 防御方均衡策略概率随资产价值V的变化趋势

图 5 攻防双方均衡策略概率随资产价值V的变化趋势(SL=2)

6 结论

本文针对侧信道安全领域缺乏系统化攻防策略选择依据的问题，构建了一个融合行为经济学与层次分析法的不完全信息静态博弈模型。核心贡献包括：(1)提出“客观成本量化—主观价值调整—工程约束”三层融合的成本结构化框架，将模糊的成本概念转化为可解释、可计算的层次化指标，克服了传统主观赋值法的随意性；(2)建立安全等级依赖的行为参数体系，将风险偏好与损失厌恶心理纳入博弈分析，使模型能够自适应不同安全场景；(3)通过贝叶斯-纳什均衡求解与敏感性分析，揭示了资产价值与防护强度之间的非线性临界关系，为“适度安全”理念提供了首个量化决策工具。

模型中的关键参数(成功率矩阵、行为参数、先验信念)目前依赖专家经验设定，虽在当前缺乏实测数据的背景下具有务实合理性，但可能影响模型的泛化性能。此外，静态博弈框架未刻画攻防双方的长期策略演化过程。未来工作将从以下方向展开：(1)数据驱动校准——基于公开侧信道攻防数据集推动参数设定的统计推断与校准；(2)模型动态化扩展——将静态博弈扩展为随机博弈或重复博弈，引入强化学习方法刻画策略演化；(3)工程化验证——与密码模块检测机构合作，在实际测评流程中检验模型的实用性与鲁棒性。

参 考 文 献

[1] KOCHER P, JAFFE J, and JUN B. Differential power analysis[C]. 19th Annual International Cryptology Conference on Advances in Cryptology, Santa Barbara, USA, 1999: 388–397. doi: 10.1007/3-540-48405-1_25.

[2] QUISQUATER J J and SAMYDE D. ElectroMagnetic analysis (EMA): Measures and counter-measures for smart cards[C]. Smart Card Programming and Security International Conference on Research in Smart Cards, Cannes, France, 2001: 200–210. doi: 10.1007/3-540-45418-7_17.

[3] KOCHER P C. Timing attacks on implementations of

Diffie-Hellman, RSA, DSS, and other systems[C]. 16th Annual International Cryptology Conference on Advances in Cryptology, Santa Barbara, USA, 1996: 104–113. doi: 10.1007/3-540-68697-5_9.

[4] ISO. ISO/IEC 17825: 2024 Information technology—security techniques—testing methods for the mitigation of non-invasive attack classes against cryptographic modules[S]. Geneva: ISO, 2024.

[5] YD/T 6305-2024. 数字内容保护技术要求[S]. 北京: 邮电部标准出版社, 2024. **(查阅网上资料, 未找到本条文献信息, 请确认)**.

[6] 姚剑波, 张涛. 基于互信息博弈的侧信道攻击安全风险评估[J]. 计算机科学, 2012, 39(S1): 69–71.
YAO Jianbo and ZHANG Tao. Side channel risk evaluation based on mutual information game[J]. *Computer Science*, 2012, 39(S1): 69–71.

[7] WANG Bing, LOU Wenjing, and HOU Y T. Modeling the side-channel attacks in data deduplication with game theory[C]. 2015 IEEE Conference on Communications and Network Security (CNS), Florence, Italy, 2015: 200–208. doi: 10.1109/CNS.2015.7346829.

[8] GENG Hui, KWIAT K A, KAMHOUA C A, *et al*. On random dynamic voltage scaling for internet-of-things: A game-theoretic approach[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2018, 37(1): 123–132. doi: 10.1109/TCAD.2017.2717780.

[9] HALPERN J and TEAGUE V. Rational secret sharing and multiparty computation: Extended abstract[C]. Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing, Chicago, USA, 2004: 623–632. doi: 10.1145/1007352.1007447.

[10] 彭长根, 田有亮, 刘海, 等. 密码学与博弈论的交叉研究综述[J]. 密码学报, 2017, 4(1): 1–15. doi: 10.13868/j.cnki.jcr.000158.
PENG Changgen, TIAN Youliang, LIU Hai, *et al*. A survey on the intersection of cryptography and game theory[J]. *Journal of Cryptologic Research*, 2017, 4(1): 1–15. doi: 10.13868/j.cnki.jcr.000158.

[11] 刘小虎, 张恒巍, 张玉臣, 等. 基于博弈论的网络攻防行为建模与态势演化分析[J]. 电子与信息学报, 2021, 43(12): 3629–3638. doi: 10.11999/JEIT200628.
LIU Xiaohu, ZHANG Hengwei, ZHANG Yuchen, *et al*. Modeling of network attack and defense behavior and analysis of situation evolution based on game theory[J]. *Journal of Electronics & Information Technology*, 2021, 43(12): 3629–3638. doi: 10.11999/JEIT200628.

[12] SUN Haiyan, SHAO Chenglong, ZHANG Jianwei, *et al*. Evolution analysis of network attack and defense situation based on game theory[J]. *Computers, Materials and Continua*, 2025, 83(1): 1451–1470. doi: 10.32604/cmc.2025.059724.

[13] BAI Xue and BAI Yongguo. Equilibrium strategy of attack and defense in computer networks based on Markov signal game theory[J]. *Journal of Cyber Security and Mobility*, 2025, 14(1): 127–153. doi: 10.13052/jcsm2245-1439.1416.

[14] ZHENG Tianshuai, DU Ye, HUA Kaisheng, *et al*. Predictive analytics for cyber-attack timing in power Internet of Things: A FlipIt game-theoretic approach[J]. *Internet of Things*, 2025, 30: 101522. doi: 10.1016/j.iot.2025.101522.

[15] ZHU Zinan and ZHOU Lin. Application of complex network attack and defense time game model in network security defense decision[J]. *Journal of Cyber Security and Mobility*, 2025, 14(2): 311–337. doi: 10.13052/jcsm2245-1439.1423.

[16] HE Long, SUN Geng, Niyato D, *et al*. Generative AI for game theory-based mobile networking[J]. *IEEE Wireless Communications*, 2025, 32(1): 122–130. doi: 10.1109/MWC.007.2400133.

[17] MAO Shaoguang, CAI Yuzhe, XIA Yan, *et al*. ALYMPICS: LLM agents meet game theory[C]. Proceedings of the 31st International Conference on Computational Linguistics, Abu Dhabi, UAE, 2025: 2845–2866.

[18] TARIQ A, SALLABI F, SERHANI M A, *et al*. Leveraging game theory and XAI for data quality-driven sample and client selection in trustworthy split federated learning[J]. *IEEE Transactions on Consumer Electronics*, 2025, 71(2): 6686–6699. doi: 10.1109/TCE.2025.3543209.

[19] PIRHOSSEINI H, ZAREI E, REZVANI M H, *et al*. Game-theoretic methods for edge/fog/cloud computation offloading: A systematic review[J]. *Computing*, 2025, 107(8): 166. doi: 10.1007/s00607-025-01515-x.

[20] 张鸿, 廖彧歆, 王汝言, 等. 面向密集场景的空天地网络资源分配算法[J]. 电子与信息学报, 2024, 46(5): 1968–1976. doi: 10.11999/JEIT231086.
ZHANG Hong, LIAO Yuxin, WANG Ruyan, *et al*. Resource allocation algorithm of space-air-ground integrated network for dense scenarios[J]. *Journal of Electronics & Information Technology*, 2024, 46(5): 1968–1976. doi: 10.11999/JEIT231086.

[21] 冯·诺伊曼, 摩根斯特恩, 王文玉, 王宇, 译. 博弈论与经济行为[M]. 北京: 三联书店, 2004.
VON NEUMANN J, MORGENSTERN O, WANG Wenyu, WANG Yu, translation. Theory of Games and Economic Behavior[M]. Beijing: SDX Joint Publishing Company, 2004. **(查阅网上资料, 未找到本条文献的页码信息, 请补充)**.

[22] SAATY T L. A scaling method for priorities in hierarchical structures[J]. *Journal of Mathematical Psychology*, 1977, 15(3): 234–281. doi: 10.1016/0022-2496(77)90033-5.

[23] SAATY R W. The analytic hierarchy process—what it is and how it is used[J]. *Mathematical Modelling*, 1987, 9(3/5): 161–176. doi: 10.1016/0270-0255(87)90473-8.

[24] KAHNEMAN D and TVERSKY A. Prospect theory: An analysis of decision under risk[M]. MacLean L C and Ziemba W T. World Scientific Handbook in Financial Economics Series: Handbook of the Fundamentals of Financial Decision Making, 2013: 99–127. doi: 10.1142/9789814417358_0006.**(查阅网上资料,未找到本条文献的出版地和出版者信息,请确认)**.

[25] OŻGA S A. Aspects of the theory of risk-bearing[J]. *Economica*, 1966, 33(130): 251–252. doi: 10.2307/2552628.

[26] 国家密码管理局. GM/T 0083-2020 密码模块非入侵式攻击缓解技术指南[S]. 北京: 中国标准出版社, 2021.
State Cryptography Administration. GM/T 0083-2020 Guideline for the mitigation of non-invasive attacks against cryptographic modules[S]. Beijing: Standards Press of China, 2021. **(查阅网上资料, 未找到本条文献作者英文信息, 请确认)**.

蔡爵嵩：男，博士生，研究方向为侧信道分析与防护.
严迎建：男，教授，研究方向为芯片安全.
王晋东：男，教授，研究方向为系统安全.

责任编辑：廖海贝

A Behavioral Economics-Based Game Model for Side-Channel Security Attack and Defense Strategies

CAI Juesong　　YAN Yingjian　　WANG Jindong

(*Information Engineering University*, *Zhengzhou* 450001, *China*)

Abstract:

Objective The field of side-channel security currently lacks a systematic, quantifiable, and reproducible model for guiding the selection of attack and defense strategies, particularly in real-world engineering contexts where resource constraints necessitate informed cost-benefit trade-offs. The absence of such a framework impedes the practical realization of the "appropriate security" principle, often leading to either over-protection or under-protection of cryptographic modules. Traditional approaches to evaluating attack and defense costs rely heavily on subjective expert judgments, which are inherently arbitrary, difficult to replicate, and lack a structured multi-dimensional assessment. To bridge this critical gap, this research proposes an interdisciplinary model that integrates game theory, behavioral economics, and the Analytic Hierarchy Process (AHP). The primary objective is to establish a holistic decision-support system that not only quantifies the multi-faceted costs of various side-channel strategies but also incorporates the psychological dimensions of decision-making under risk, thereby enabling dynamic and economically rational security strategy selection tailored to specific asset values and security levels.

Methods This study constructs a multi-layered modeling framework based on a static non-cooperative game with incomplete information. First, the side-channel analyst and the defense designer are formally defined as rational players, each possessing a finite set of strategies: the attacker may choose from non-modeling attacks such as DPA/CPA, modeling-based attacks like template attacks, or emerging deep learning-based side-channel analysis; the defender may adopt countermeasures including time hiding, amplitude hiding, or masking/blinding techniques. To systematically quantify the often-overlooked cost dimension, an AHP-based structured cost model is introduced. Through pairwise comparison matrices, the model decomposes costs into multiple criteria—such as time, data storage, computational resources, expertise, and hardware overhead—and assigns objective weights to each criterion, thereby replacing subjective cost estimates with a reproducible, hierarchical evaluation system. Furthermore, to reflect real-world decision-making behavior, key concepts from behavioral economics are integrated: Prospect Theory models how gains and losses are perceived relative to a reference point, while risk aversion coefficients capture players' tolerance for uncertainty. These behavioral parameters are explicitly linked to the security level of the cryptographic module, allowing the model to adapt to different operational contexts. The resulting behavioral-augmented Bayesian game is then solved using the concept of Bayes-Nash Equilibrium, wherein each player's optimal mixed strategy is derived based on their private type

(behavioral profile) and beliefs about the opponent. To ensure engineering relevance, the As Low As Reasonably Practicable principle is incorporated as a constraint, enforcing that any selected defense strategy must be justifiable in terms of risk reduction versus cost incurred. Numerical solutions are obtained via a customized sequential quadratic programming algorithm implemented in Python.

Results and Discussions A comprehensive experimental evaluation was conducted to validate the proposed model's consistency, sensitivity, and practical utility. The AHP-based cost quantification demonstrated strong internal consistency, with all consistency ratios below the 0.1 threshold, confirming the reliability of the judgment matrices. The derived weight distributions revealed intuitive priorities: attackers placed greater emphasis on technical barriers and computational cost, whereas defenders prioritized design complexity and performance overhead. The behavioral adjustment layer successfully modulated perceived costs according to security levels: under low-security conditions (high risk aversion), costs were perceptually inflated, leading to conservative strategy choices; under high-security conditions, decision-making aligned more closely with objectively quantified costs. Equilibrium analysis across varying asset values and security levels yielded interpretable and rational strategy profiles. For low-value assets, both players exhibited a strong tendency toward low-cost or "no action" strategies, adhering to the lower bound of the ALARP region. As asset value increased, a clear threshold effect was observed, triggering a shift toward high-cost, high-efficacy strategies such as deep learning-based attacks and masking-based defenses. Sensitivity analysis further confirmed that defense strategy probabilities increased monotonically with asset value, validating the model's ability to capture the non-linear relationship between protection intensity and asset criticality. These findings underscore the model's capacity to support context-aware, adaptive security decision-making that balances risk, cost, and psychological factors.

Conclusions This research presents a novel, behaviorally informed game-theoretic model for side-channel security strategy selection, addressing a significant void in existing literature regarding structured cost-benefit assessment. By integrating AHP-based objective cost quantification, behaviorally adjusted subjective valuations, and ALARP-driven engineering constraints, the proposed framework offers a multi-dimensional, reproducible, and context-sensitive tool for analyzing attack-defense interactions. The model advances the field by explicitly linking security levels to behavioral parameters, enabling dynamic strategy adaptation in response to both asset value and decision-makers' risk perceptions. Although the current implementation relies partially on expert-defined parameters and operates within a static game setting, it establishes a critical foundation for transitioning from heuristic-based security decisions to quantitatively grounded, interdisciplinary analysis. Future work will focus on parameter calibration using real-world attack/defense datasets, extension to multi-stage dynamic games to capture strategic evolution over time, and empirical validation in industrial cryptographic evaluation scenarios. This study contributes the first systematic methodology for cost-aware, behaviorally realistic strategy optimization in side-channel security, offering both theoretical insights and practical guidance toward achieving "appropriate security" in cryptographic engineering.

Key words: Side-channel security; Attack and defense strategies; Behavioral economics; Analytic Hierarchy Process (AHP); Static game with incomplete information