



面向高级持续性威胁的卫星互联网最优修复策略

王云涛, 何春林, 苏洲*

西安交通大学网络空间安全学院, 西安 710049

* 通信作者. E-mail: zhousu@xjtu.edu.cn

收稿日期: 2025-11-01; 修回日期: 2026-01-01; 接受日期: 2026-02-05; 网络出版日期: 2026-02-28

国家自然科学基金 (批准号: 62302387, U24A20237, U22A2029) 资助项目

摘要 随着低轨卫星星座与空间信息基础设施的快速发展, 卫星互联网已成为全球信息网络的重要组成部分. 然而, 其链路时变性、星地异构性与节点资源受限等特性使得在面对高级持续性威胁时呈现出高隐蔽性与防御复杂性, 传统基于博弈论的静态或半静态拓扑下的防御策略难以适应卫星网络的高动态环境. 针对该问题, 本文面向时变卫星互联网提出了一种高效动态的最优安全修复策略. 首先, 构建动态序贯修复博弈模型, 将攻防过程建模为有限时域上的多时隙序贯零和子博弈来刻画攻击者的渐进式渗透行为与防御者的动态修复决策. 其次, 考虑拓扑动态演化、节点状态差异、防御资源约束等因素, 提出基于最优控制的前后向搜索与策略更新算法, 以低计算复杂度求解博弈均衡策略. 仿真结果表明, 所提方法能在资源受限条件下有效降低攻击累积损害并提升防御鲁棒性.

关键词 卫星互联网安全, 高级持续性威胁, 动态博弈, 最优修复

1 引言

近年来, 随着低轨 (low Earth orbit, LEO) 卫星星座、大型通信星座及空间信息基础设施的快速部署, 卫星互联网 (satellite Internet) 逐步成为地面互联网的重要延伸与全球覆盖的关键通道^[1,2]. 相比传统地面网络, 卫星互联网在广域覆盖、快速投送及抗灾恢复等方面具有显著优势, 但其独特的物理与运行特性亦带来了新的安全威胁, 尤其是高级持续性威胁 (advanced persistent threat, APT) 风险持续上升^[3]. 由于链路时变性与频繁切换导致的连通性瞬时变化、往返时延及链路质量的时空差异, 星间链路 (inter-satellite link, ISL) 与地面链路的异构特性、节点资源的高度受限, 以及遮挡与跃迁现象造成的检测与响应延迟, 使得卫星网络在防御层面呈现出高隐蔽性、低可观测性与安全响应慢等特征, 从而成为 APT 攻击者重点利用的高价值目标^[4]. 因此, 针对卫星互联网的安全建模、快速响应与资源受限下的最优修复策略, 已成为亟需解决的研究问题.

在应对 APT 与复杂对抗行为的研究中, 博弈论因其能够形式化刻画攻防双方的策略交互与收益权衡, 成为近年来网络安全建模的重要理论工具^[5~15]. 现有研究已从防御部署优化^[5~7]、入侵检测策略^[8,9]、恶意软件

引用格式: 王云涛, 何春林, 苏洲. 面向高级持续性威胁的卫星互联网最优修复策略. 中国科学: 信息科学, 2026, 56: 603–619, doi: 10.1360/SSI-2025-0462

Wang Y T, He C L, Su Z. Optimal recovery strategies for satellite Internet under advanced persistent threats. Sci Sin Inform, 2026, 56: 603–619, doi: 10.1360/SSI-2025-0462

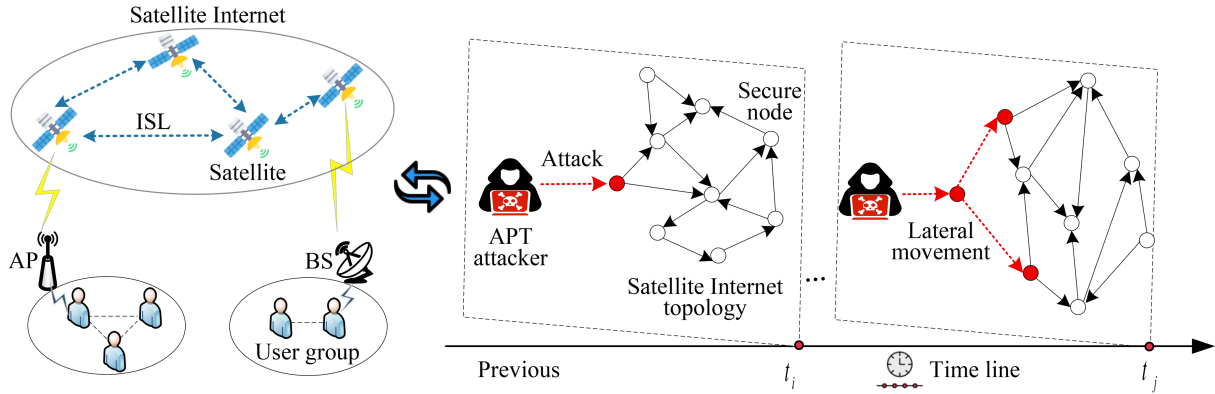


图 1 (网络版彩图) 卫星互联网拓扑时变下的 APT 横向移动攻击示意图.

Figure 1 (Color online) APT lateral propagation over a time-varying satellite Internet topology.

传播抑制^[10~13]及系统修复与恢复策略^[14,15]等角度构建了多种博弈模型,为量化攻防行为及收益关系提供了坚实的理论基础.例如,Zhang等人^[5]揭示了防御方频繁策略调整可能造成关键信息泄露,构建了动态APT对抗博弈模型,并基于信息泄露最小化原理推导出最优策略调整时机的纳什均衡解.Shen等人^[13]针对工业物联网环境下的恶意软件传播风险,提出了基于微分博弈的双层传播-修复模型,结合主观努力参数与优化理论求解设备层与中心层的最优防控策略,并引入深度强化学习算法设计了传播动态抑制方法.Bi等人^[15]则基于Stackelberg博弈提出动态攻防策略,将节点状态划分为安全与不安全两类,并利用遗传算法实现快速均衡求解.然而,将上述攻防博弈策略直接迁移至卫星互联网场景仍面临挑战.大多数现有研究隐含或显式地假设网络拓扑在防御周期内相对稳定,而卫星网络的轨道运动与链路可用性导致拓扑在短时间尺度上频繁变化,这使得传统静态拓扑下博弈模型和离线优化策略难以满足卫星互联网实际部署需求.

近年来,国内外学者围绕卫星互联网安全问题,陆续提出了多种基于博弈论的干扰对抗机制^[16~19]与路由优化策略^[20].然而,现有面向卫星互联网的安全博弈研究很少考虑APT攻击动态横向移动特性的影响,如图1所示,网络拓扑的时变性的会导致攻击者横向移动路径更复杂、攻击面瞬时变化更频繁,从而要求防御端在资源受限条件下实现在线的低延迟的策略更新与执行.实际案例亦表明,动态拓扑可被攻击者利用以规避传统静态防护,例如2023年DEF CON黑客大会Hack-A-Sat 4竞赛结果表明^[21],即便在高安全等级和受限通信条件下运行的真实在轨卫星,仍可能因软硬件耦合漏洞、协议不完备或权限管理缺陷而遭受入侵,表明在轨卫星网络的攻防边界正从地面控制中心延伸至星间链路层,卫星互联网面临更复杂的横向移动安全威胁.此外,攻防信息往往存在不完备性与时延性,如遥测误差、传输延迟或跨域日志缺失,进一步提升了策略自适应性与鲁棒性设计的难度.综上,面向时变卫星网络的APT防护机制需兼顾拓扑变化性、节点状态演变性、攻防信息不确定性以及防御资源动态限制.

为应对上述挑战,本文提出一种面向时变卫星互联网的动态修复博弈建模与高效求解方法.主要贡献包括:(1)提出了针对卫星互联网的节点级安全状态演化模型和序贯修复博弈模型,将APT攻击过程刻画为有限时域上的多时隙序贯子博弈序列,通过在总体动态博弈中嵌入子博弈结构,形式化描述了防御者在时变拓扑与节点状态演化下的动态修复决策过程与攻击者的渐进式渗透行为;(2)考虑到拓扑时变性、威胁影响差异、节点状态异质性及防御资源的动态约束,提出了动态拓扑下的节点重要性评估算法,设计了一种基于最优控制的前后向搜索与策略更新方法,该算法能够在每一时隙内以低计算成本逼近博弈均衡,生成近最优可执行的动态修复策略,实现攻防双方的策略自适应优化;(3)在多种攻击能力配置与网络动态场景下进行大量仿真评估,结果表明,所提方法在降低攻击损害和提升系统恢复效率方面显著优于基线方案,并能在资源受限条件下实现高效防御与自适应优化.

2 相关工作

2.1 卫星互联网安全防护机制

随着卫星互联网的快速发展,其安全问题受到日益广泛的关注.文献[22]提出了一种基于完全量子纠缠的星地通信方法,在无需依赖中继设备的条件下将卫星互联网地面终端的安全通信距离提升约10倍,显著增强了卫星通信过程的安全性.文献[23]设计了一种基于节点信任评估的安全路由算法,利用分布式信任计算模型对卫星节点进行实时信任值评估,从而检测并隔离恶意节点,缓解了内部攻击对网络稳定性的破坏.为进一步优化资源调度与安全管理,文献[24]系统综述了人工智能在大规模卫星网络中的应用进展,涵盖信道预测、频谱感知、资源分配、星间链路优化及安全防护等方面,指出人工智能可显著增强卫星网络的自适应性及抗干扰性.文献[25]提出了一种基于联合异常检测的卫星安全传输优化算法,通过将异常检测机制与传输控制相结合并使用最优控制方法,实现了卫星网络中潜在恶意行为的识别与动态资源优化,有效提升了数据传输安全性与网络服务稳定性.此外,针对星地融合内容分发问题,文献[26]提出了一种星地融合网络的双层协同缓存与内容分发优化模型,在卫星与地面节点之间实现多层缓存协作与传输优化,有效缓解了星间链路的带宽压力并提升了数据服务可靠性.

尽管上述研究在星地通信安全、节点信任管理及资源调度优化等方面取得了显著进展,但现有工作普遍侧重于单周期的防御建模,较少关注卫星网络中节点安全状态的动态演化特征与APT攻击的长期隐蔽性问题.

2.2 卫星互联网攻防博弈机制

近年来,基于博弈论的网络威胁防御机制设计已成为企业网与物联网安全的研究热点.文献[7]提出了一种基于多智能体深度强化学习的微分博弈欺骗资产选择算法,弥补了传统方法在攻防建模与实时决策方面的不足.文献[10]针对多攻击者-多防御者场景,构建了平均场博弈模型以实现防御者间的协同响应,从而抑制恶意软件传播.文献[11]基于信号博弈设计了一种主动入侵检测机制,通过贝叶斯更新动态调整检测策略,以提升检测精度与对抗适应性.文献[12]漏洞依赖图与Stackelberg攻防博弈模型,在企业级网络环境下实现漏洞修复与防御策略优化.文献[9]提出了一种基于随机博弈的动态信息流追踪框架,为检测与防御APT攻击提供了新思路.

同时,博弈论方法在具有开放信道与动态拓扑特征的卫星互联网安全防护中也受到持续关注.文献[16]提出了基于深度强化学习与Stackelberg博弈的空间抗干扰机制,通过分层博弈建模用户与智能干扰源的对抗交互,采用深度强化学习获取可行路由子集并结合Q学习实现快速响应,有效提升了异构卫星网络的抗干扰能力与路由效率.文献[20]构建了基于极大极小值理论的卫星攻防模型,考虑攻击者在简单攻击和复杂攻击两种策略下,分别通过固定节点攻击和概率化攻击序列以最大化传输时延,并基于路由表搜索实现防御路径优化.文献[17]提出了一种分层抗干扰Stackelberg博弈框架,来建模干扰器与卫星终端设备间的对抗交互,并利用联盟形成博弈与Q学习实现分布式动态抗干扰子网构建,从而在能耗与抗干扰性能之间实现平衡.文献[18]将用户调度与功率分配的联合优化建模为势博弈,理论证明了纳什均衡的存在性并设计低复杂度迭代算法,有效实现了同信道干扰规避与系统容量提升.文献[19]针对卫星-无人机一体化系统中的全双工窃听与干扰攻击,提出了一种双层Stackelberg博弈模型并设计了三阶段最优响应迭代算法以求解博弈均衡,通过源无人机与攻击者间的功率博弈以及无人机网络内的协作激励机制,实现了节点协作下的安全传输.

然而,现有基于博弈论的卫星互联网安全研究多聚焦于抗干扰博弈、路由优化、功率优化、静态拓扑与简单节点状态下的防御部署优化,尚未充分考虑APT攻击的动态渗透特性与卫星网络拓扑的时变性.相比之下,本文针对APT攻击所特有的长期潜伏、渐进渗透与横向移动特性,构建了节点安全状态多阶段演化模型(S-L-A-Q-H),并将攻防交互嵌入到连续时间的状态动力学中;提出了动态序贯修复博弈,将整体攻防过程分解为多时隙子博弈,在每一时隙内形成“攻击先行-防御响应-状态更新”的闭环反馈;设计了综合考虑APT横向传播、节点安全状态演化与资源约束下的最优修复决策算法,能以较低计算复杂度在线逼近博弈均衡解.

3 系统模型

3.1 网络模型

设图 $\mathbb{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ 表示在时刻 t 所考虑范围内卫星互联网中 LEO 卫星之间的网络拓扑连接关系, 其中 $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ 表示 LEO 卫星集合, $\mathcal{E}(t)$ 表示在时刻 t 存在的卫星间通信链路集合. 边 $\{v_m, v_n\} \in \mathcal{E}(t)$ 表示 LEO 卫星 v_m 与 v_n 之间存在直接通信链路. 令 $A(t) = [a_{mn}(t)]_{N \times N}$ 为图 $\mathbb{G}(t)$ 的邻接矩阵, 其中 $a_{mn}(t) = \{0, 1\}$. 若 $\{v_m, v_n\} \in \mathcal{E}(t)$, 则 $a_{mn}(t) = 1$, 否则, $a_{mn}(t) = 0$.

如图 1 所示, 假设所考虑的卫星网络运营者 (以下简称“防御者”) 在时刻 $t = 0$ 侦测到安全异常事件, 并在时间窗口 $[0, T]$ 内实施一系列“检测 – 隔离 – 加固” (detect-quarantine-harden, DQH) 安全操作以遏制威胁传播与系统失陷. 在任意时刻 $t \in [0, T]$, 网络中的每个卫星节点 v_n 可依据其在安全态势下的行为特征进一步划分为下述 5 类互斥状态: 正常 (susceptible, S)、潜伏 (latent, L)、激活 (active, A)、隔离 (quarantined, Q) 与加固 (hardened, H). 其中, 潜伏态和激活态统称为失陷态 (compromised, C). 5 种状态的定义如下.

- 正常状态 S : 卫星处于正常运行状态, 未被攻陷, 无已知后门或异常活动, 正常提供通信服务.
- 潜伏状态 L : 卫星已被攻击者渗透或植入隐蔽载荷, 但尚未表现出明显恶意行为, 此阶段常伴随恶意指令的缓存或信息的隐性收集.
- 激活状态 A : 卫星已被完全攻陷, 攻击者正在利用该节点实施指令与控制、横向移动、数据窃取、链路阻塞或信息篡改等恶意行为.
- 隔离状态 Q : 卫星被防御者检测到存在异常迹象后被暂时隔离, 断开通信连接以阻止威胁扩散.
- 加固状态 H : 卫星经过系统恢复、策略更新与补丁修复后重新投入运行, 对当前攻击具有一定免疫性, 但其加固效力会随时间衰减, 最终可能回到正常态.

定义1 (卫星安全状态变量) 定义卫星 v_n 在时刻 t 的安全状态指示变量为 $\beta_n(t) \in \{0, 1, 2, 3, 4\}$, 表示分别处于正常状态 S 、潜伏状态 L 、激活状态 A 、隔离状态 Q 与加固状态 H . 卫星 v_n 在时刻 t 的安全状态向量定义为

$$x_n(t) = [S_n(t) \ L_n(t) \ A_n(t) \ Q_n(t) \ H_n(t)]^\top, \quad (1)$$

其中, $S_n(t), L_n(t), A_n(t), Q_n(t), H_n(t) \in [0, 1]$ 分别表示卫星 v_n 在时刻 t 处于正常态、潜伏态、激活态、隔离态与加固态的概率, 即 $S_n(t) = \Pr\{\beta_n(t) = 0\}$, $L_n(t) = \Pr\{\beta_n(t) = 1\}$, $A_n(t) = \Pr\{\beta_n(t) = 2\}$, $Q_n(t) = \Pr\{\beta_n(t) = 3\}$, $H_n(t) = \Pr\{\beta_n(t) = 4\}$. 并满足归一化约束: $S_n(t) + L_n(t) + A_n(t) + Q_n(t) + H_n(t) = 1$. 进而, 整个卫星网络在时刻 t 的安全状态可表示为向量

$$X(t) = (x_1(t), x_2(t), \dots, x_N(t)). \quad (2)$$

在实际运行中, 可通过文献 [27, 28] 中所提的 APT 检测程序获取初始状态 $X(0)$.

3.2 攻击渗透与横向移动模型

对于任意处于未失陷 (即正常) 状态的卫星节点 v_n , 其被攻陷的途径可分为直接渗透 (direct penetration) 与横向移动 (lateral movement) 两类^[15], 这两种机制分别对应攻击者通过外部入侵和内部传播实现对目标卫星节点控制的不同过程.

(1) 直接渗透. 令 $\lambda_n^{\text{dir}}(t)$ 表示时刻 t 节点 v_n 的直接渗透率, 其大小与攻击者对该节点施加的攻击强度正相关. 设 $u_n(t)$ 为作用于节点 v_n 的攻击控制变量, 用以刻画攻击者在时刻 t 内的渗透或攻击资源占用策略. 则有

$$\lambda_n^{\text{dir}}(t) = \lambda_0^{\text{dir}} u_n(t), \quad (3)$$

其中, $\lambda_0^{\text{dir}} > 0$ 表示单位攻击强度下的基准渗透能力.

(2) 横向移动. 在卫星互联网中, 攻击者除直接渗透外, 还可通过受控节点在网络内部进行横向移动攻击. 横向传播速率与源节点的攻击强度、链路可达性及源节点处于失陷状态的概率等因素正相关. 为此, 定义由节点

v_m 向节点 v_n 的横向传播率为

$$\lambda_{mn}^{\text{lat}}(t) = \lambda_0^{\text{lat}} u_n(t) R_{mn}(t), \quad (4)$$

其中, $\lambda_0^{\text{lat}} > 0$ 为横向传播基准系数, $u_n(t)$ 为源节点 v_n 遭受的攻击强度, $R_{mn}(t) \in [0, 1]$ 为从 v_m 到 v_n 的星间链路瞬时传输速率. 综合考虑星间链路连通性、可用带宽、路径损耗与并发干扰等因素, 根据香农公式, 对任意星间链路 $\{v_m, v_n\} \in \mathcal{E}(t)$, 其瞬时传输速率表示为^[29, 30]

$$R_{mn}(t) = W_{mn} \log_2 \left(1 + \frac{P_m G_{mn} h_{mn}^f(t) h_{mn}^a}{\varkappa_{mn}(t) + k_0 T_0 W_{mn}} \right), \quad (5)$$

其中, W_{mn} 为 ISL 链路 $\{v_m, v_n\}$ 的带宽; P_m 为卫星 v_m 的发射功率; G_{mn} 为收发天线的联合增益; $h_{mn}^f(t)$ 为自由空间损耗, 其值随几何距离平方反比衰减^[30]

$$h_{mn}^f(t) = \left(\frac{c_0}{4\pi d_{mn}(t) f_{mn}^c} \right)^2, \quad (6)$$

其中, c 为光速, $d_{mn}(t)$ 为两星在时刻 t 的几何距离, f_{mn}^c 为中心频率; h_{mn}^a 为附加线路损耗, 如方向性损耗、连接器损耗等; k_0 为玻尔兹曼常数; T_0 为系统噪声温度. 星间链路 $\{v_m, v_n\} \in \mathcal{E}(t)$ 所受的并发干扰 $\varkappa_{mn}(t)$ 可由网络内并发链路的叠加效应刻画为

$$\varkappa_{mn}(t) = \sum_{\{v_i, v_j\} \in \Lambda_{mn}(t)} y_{ij}(t) P_i G_{ij} h_{ij}^f(t) h_{ij}^a, \quad (7)$$

其中, $\Lambda_{mn}(t) \subseteq \mathcal{E}(t)$ 为对链路 $\{v_m, v_n\}$ 构成干扰的 ISL 链路集合; $y_{ij}(t) \in \{0, 1\}$ 为激活指示变量; $y_{ij}(t) = 1$ 表示该链路 $\{v_i, v_j\}$ 在时刻 t 正处于传输状态且会产生干扰, $y_{ij}(t) = 0$ 表示静默状态.

节点 v_n 因横向移动被感染的总速率可表示为所有邻接节点传播效应的加权叠加

$$\lambda_n^{\text{lat}}(t) = \sum_{m=1, m \neq n}^N a_{mn} \lambda_{mn}^{\text{lat}}(t). \quad (8)$$

综合直接渗透与横向传播两种模式, 卫星 v_n 的总体攻击渗透率可表示为

$$\lambda_n(t) = \lambda_n^{\text{dir}}(t) + \lambda_n^{\text{lat}}(t) = \lambda_0^{\text{dir}} u_n(t) + \sum_{m=1, m \neq n}^N a_{mn} \lambda_0^{\text{lat}} u_n(t) R_{mn}(t). \quad (9)$$

(3) 攻防控制变量. 状态向量 $X(t)$ 的动态演化由攻击者与防御者的联合决策所驱动.

- **攻击者控制变量.** 攻击者的渗透策略向量定义为

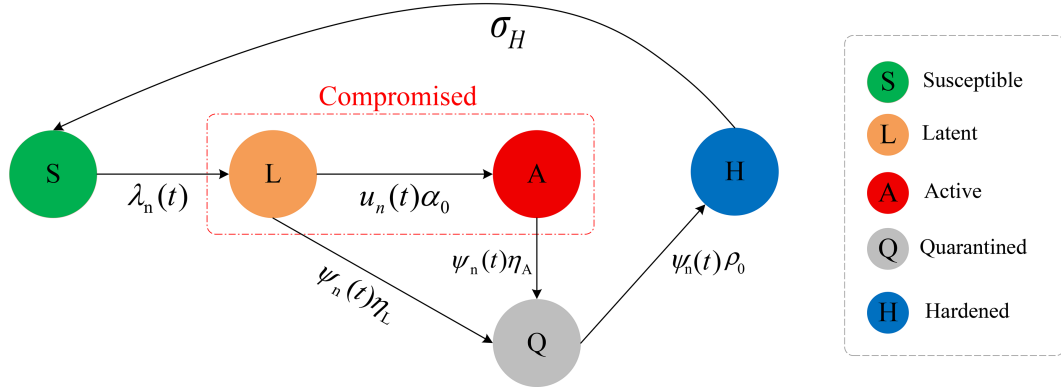
$$u(t) = (u_1(t), u_2(t), \dots, u_N(t)), \quad 0 \leq u_n(t) \leq u_a^{\max}, \quad \forall n \in \{1, \dots, N\}, \quad (10)$$

其中, $u_n(t)$ 表示攻击者在时刻 t 对卫星节点 v_n 的攻击强度策略. $u_a^{\max}$ 为单节点可分配的最大攻击强度或攻击资源上限. $u_n(t) = 0$ 表示该卫星节点当前未被攻击.

- **防御者控制变量.** 防御者的修复策略向量定义为

$$\psi(t) = (\psi_1(t), \psi_2(t), \dots, \psi_N(t)), \quad 0 \leq \psi_n(t) \leq \psi_d^{\max}, \quad \forall n \in \{1, \dots, N\}, \quad (11)$$

其中 $\psi_n(t)$ 表示防御者在时刻 t 对卫星 v_n 采取的防御强度, 包括检测、隔离与加固等操作. $\psi_d^{\max}$ 表示防御者在单节点可施加的最大防御强度. $\psi_n(t) = 0$ 表示卫星 v_n 处于无防护状态.

图2 (网络版彩图) 时刻 t 卫星 v_n 安全状态转移图.Figure 2 (Color online) Security state transition of satellite v_n at time t .

3.3 卫星安全状态演变模型

卫星互联网系统的安全状态随时间动态演化, 可通过一组非线性微分方程进行刻画. 系统中每个卫星节点在不同安全阶段之间的状态迁移受到攻击者渗透强度、防御者修复策略以及节点间拓扑关系的共同影响. 如图2所示, 可建立如下非线性动力学传播模型.

定理1 时刻 t 卫星互联网中任意卫星 $v_n, \forall n \in \mathcal{N}$ 的安全状态演化由下列微分方程组描述:

$$\dot{S}_n(t) = - \left[\lambda_n^{\text{dir}}(t) + \sum_{m=1, m \neq n}^N \lambda_{mn}^{\text{lat}}(t) a_{mn} \right] S_n(t) + \sigma_H H_n(t), \quad (12)$$

$$\dot{L}_n(t) = \left[\lambda_n^{\text{dir}}(t) + \sum_{m=1, m \neq n}^N \lambda_{mn}^{\text{lat}}(t) a_{mn} \right] S_n(t) - \alpha_0 u_n(t) L_n(t) - \psi_n(t) \eta_L L_n(t), \quad (13)$$

$$\dot{A}_n(t) = \alpha_0 u_n(t) L_n(t) - \psi_n(t) \eta_A A_n(t), \quad (14)$$

$$\dot{Q}_n(t) = \psi_n(t) [\eta_L L_n(t) + \eta_A A_n(t) - \rho_0 Q_n(t)], \quad (15)$$

$$\dot{H}_n(t) = \rho_0 \psi_n(t) Q_n(t) - \sigma_H H_n(t), \quad (16)$$

其中, α_0 为潜伏转化率, 即潜伏态节点在攻击强度 $u_n(t)$ 下转化为激活态的速率; $\lambda_n^{\text{dir}}(t)$ 与 $\lambda_{mn}^{\text{lat}}(t)$ 分别表示卫星节点 v_n 的直接渗透率与来自邻接节点 v_m 横向传播的间接渗透率; a_{mn} 为网络邻接矩阵元素; η_L 与 η_A 分别表示防御者对潜伏态和激活态节点的检测与隔离效率; ρ_0 表示被隔离态节点恢复至安全状态的速率; σ_H 表示加固态节点的自然失效衰减率.

证明 证明略, 可参考现有文献 [14, 15].

4 面向卫星互联网的动态序贯修复博弈框架

4.1 攻防成本函数

(1) 卫星互联网中攻击者的成本函数定义为

$$C_A(t) = \sum_{n=1}^N r_a(u_n(t))^2 + c_L V_n L_n(t) + c_A V_n A_n(t), \quad (17)$$

其中, $u_n(t)$ 为攻击者在节点 v_n 上施加的渗透强度, c_L 与 c_A 分别为潜伏态和激活态节点的维持成本系数, V_n 表示卫星 n 的重要性得分 (具体见 5.2 小节). 第一项 $r_a(u_n(t))^2$ 为攻击强度成本, 采用二次型形式体现攻击资源

的边际收益递减效应^[31], 即攻击者若持续提高攻击强度, 其代价将呈平方级上升. 后两项分别表示维持潜伏态 $L_n(t)$ 与激活态 $A_n(t)$ 所需的资源消耗代价, 其中潜伏态节点需维持隐蔽控制链路及情报收集功能, 激活态节点则需保持持续干扰或伪造信号行为.

(2) 卫星互联网中防御者的成本函数定义为

$$C_D(t) = \sum_{n=1}^N r_d(\psi_n(t))^2 + c_Q V_n Q_n(t) + c_H V_n H_n(t), \quad (18)$$

其中, $\psi_n(t)$ 表示防御者在节点 v_n 上施加的检测、隔离与修复强度, c_Q 与 c_H 分别反映防御者在隔离态和加固态下的维护开销. 第一项 $r_d(\psi_n(t))^2$ 为防御强度成本, 亦采用二次型结构^[31], 以体现高强度防御措施 (例如高频检测、链路重配置或补丁更新) 带来的计算与通信负载的非线性增长. 后两项为系统维持成本: 隔离节点 $Q_n(t)$ 会降低网络可用性并导致任务中断, 造成运行损失, 而加固节点 $H_n(t)$ 的重启与安全验证亦需额外资源消耗.

4.2 卫星互联网最优化 APT 安全修复问题表述

综合攻击者与防御者的动态交互作用, 系统的总体性能指标定义为

$$\begin{aligned} U(u, \psi) &= \sum_{n=1}^N V_n \left(\gamma_L L_n(T)^2 + \gamma_A A_n(T)^2 - \gamma_S S_n(T)^2 - \gamma_H H_n(T)^2 - \gamma_Q Q_n(T)^2 \right) - \int_0^T (C_A(t) - C_D(t)) dt \\ &= \Phi(X(T)) - \int_0^T (C_A(t) - C_D(t)) dt, \end{aligned} \quad (19)$$

其中终端项定义为

$$\Phi(X(T)) = \sum_{n=1}^N V_n \left(\gamma_L L_n(T)^2 + \gamma_A A_n(T)^2 - \gamma_S S_n(T)^2 - \gamma_H H_n(T)^2 - \gamma_Q Q_n(T)^2 \right), \quad (20)$$

其中, $\gamma_L, \gamma_A, \gamma_S, \gamma_H, \gamma_Q$ 分别为终端时刻各安全状态的重要性系数. 终端项 $\Phi(X(T))$ 衡量系统在博弈结束时刻 T 的整体安全态势, 即攻击者期望潜伏态与激活态节点比例 $(L_n(T), A_n(T))$ 尽可能高, 以实现最大控制范围; 而防御者则期望系统中正常态 $S_n(T)$ 、加固态 $H_n(T)$ 与隔离态 $Q_n(T)$ 节点占比更高, 以提升网络的安全稳定性.

因此, 卫星互联网的 APT 最优修复问题可形式化表述为

$$\begin{aligned} &\min_{\psi(t)} \max_{u(t)} U(u, \psi) \\ &\text{s.t.} \quad \begin{cases} \dot{X}(t) = f(X(t), u(t), \psi(t)), \\ 0 \leq u_n(t) \leq u_a^{\max}, 0 \leq \psi_n(t) \leq \psi_d^{\max}, \\ X(0) = X_0. \end{cases} \end{aligned} \quad (21)$$

4.3 动态序贯修复博弈

在卫星互联网系统中, 攻击者与防御者围绕系统安全状态的演化进行长期动态对抗, 其整体交互过程可建模为一个零和微分博弈^[14]. 攻击者通过资源调度与渗透策略力图扩大受控节点比例, 而防御者则依托检测、隔离与修复机制尽可能降低系统风险, 以保持整体安全与任务连续性. 为进一步刻画这种长期对抗在时间维度上的策略演变与反馈机制, 本文将宏观的连续博弈过程细化为多个相互关联的时序子博弈^[32], 以反映攻击与防御在不同阶段的策略调整与状态更新. 因此, 系统在时间区间 $[0, T]$ 内的安全演化过程可建模为一个序贯决策博弈对抗框架.

(1) 博弈参与者. 记攻击者为 $\mathcal{A}$, 防御者为 $\mathcal{D}$. 在任意时刻 $t \in [0, T]$, 双方分别选择其决策变量

$$u(t) = (u_1(t), u_2(t), \dots, u_N(t)), \quad \psi(t) = (\psi_1(t), \psi_2(t), \dots, \psi_N(t)). \quad (22)$$

(2) 博弈策略空间. 攻击者与防御者的策略空间分别定义为

$$\mathcal{U} = \{u(t) : 0 \leq u_n(t) \leq u_a^{\max}, \forall n, t\}, \quad \Psi = \{\psi(t) : 0 \leq \psi_n(t) \leq \psi_d^{\max}, \forall n, t\}. \quad (23)$$

为体现序贯博弈特性, 假设攻击者先行决策 $u(t)$, 防御者在观测系统状态 $X(t)$ 并评估攻击强度后, 制定最优响应策略 $\psi(t) = \psi^*(u, X(t))$.

(3) 时序子博弈结构. 在每个时隙 $\Delta t_k = [t_k, t_{k+1}]$ 内, 攻击者首先选择渗透强度 $u^{(k)}$, 随后防御者根据反馈态势选择防护强度 $\psi^{(k)}$. 子博弈结束后, 系统状态更新为 $X(t_{k+1}) = f(X(t_k), u^{(k)}, \psi^{(k)})$, 并进入下一轮决策. 在 Stackelberg 框架下, 攻击者作为领导者 (leader) 先行决策, 其最优策略需预期防御者的响应; 防御者作为跟随者 (follower) 在给定 $u(t)$ 情况下, 解决下述最优化问题:

$$\psi^*(u) = \arg \min_{\psi \in \Psi} U(u, \psi). \quad (24)$$

将防御者的最优响应 $\psi^*(u)$ 代入后, 攻击者的最优化问题可表示为

$$u^* = \arg \max_{u \in \mathcal{U}} U(u, \psi^*(u)). \quad (25)$$

(4) 纳什均衡. 纳什均衡刻画了每个时隙下的序贯零和子博弈的解, 其定义如下.

定义2 (纳什均衡) 记 $u^* = (u_1^*(t), \dots, u_N^*(t))$ 与 $\psi^* = (\psi_1^*(t), \dots, \psi_N^*(t))$. 在零和框架下, 防御者选择 ψ 以最小化 U , 而攻击者选择 u 以最大化 U . 在每个时隙内, 若存在一对策略 (u^*, ψ^*) 满足对任意可行 (u, ψ) 均有

$$U(u, \psi^*) \leq U(u^*, \psi^*) \leq U(u^*, \psi), \quad \forall u \in \mathcal{U}, \psi \in \Psi, \quad (26)$$

则称 (u^*, ψ^*) 为卫星互联网动态序贯零和博弈的纳什均衡解.

5 三阶段分层最优策略求解

5.1 总体思路

卫星互联网的动态拓扑变化以及攻防交互的时序性使得系统优化策略易陷入局部最优, 为避免陷入局部最优陷阱并实现全局最优修复决策, 本文构建了一个三阶段分层求解框架, 包括态势感知、节点重要性评估与最优策略求解 3 个阶段.

阶段 I: 网络初始安全态势感知. 在该阶段, 防御者首先基于现有 APT 检测与异常识别方法^[27,28] 识别当前卫星互联网中的可疑节点及潜在威胁源, 从而得到系统初始状态 $X(0)$, 为后续决策提供输入.

阶段 II: 节点重要性评估 (见 5.2 小节). 在阶段 I 的基础上, 提出一种适用于动态拓扑的节点重要性评估算法. 该算法在变化的卫星网络拓扑下, 综合考虑节点的结构参数、传播敏感性与时序稳定性等因素, 对节点的重要性进行动态量化评估. 评估结果作为阶段 III 的优化输入参数, 用于确定不同节点的优先修复次序.

阶段 III: 基于最优控制的最优策略求解 (见 5.3 小节). 基于节点重要性, 利用 Pontryagin 极大值原理构建最优控制系统, 通过前向 - 后向迭代算法实现状态与伴随方程的联立求解, 得到系统在有限时域内的最优修复策略, 从而实现防御资源在时变环境下的分配优化.

5.2 动态拓扑下的节点重要性评估算法

在时变卫星网络中, 节点重要性取决于网络拓扑结构、APT 横向传播能力、网络拓扑变化速率等多因素. 为此, 设计了一种面向动态拓扑的节点重要性评估算法, 以支撑后续最优修复决策.

(1) 拓扑结构因子. 节点的拓扑特征决定了其在网络中的重要性. 本文采用 k-shell 分解算法^[33] 表征节点在网络中的拓扑影响范围, 并结合介数中心性 (betweenness centrality) 衡量其对网络连通性的贡献, 即 $C_n^{\text{bet}}(t) =$

$\sum_{l \neq n \neq m \in \mathcal{N}} \frac{\sigma_{lm}(n)}{\sigma_{lm}}$, 其中, σ_{lm} 表示卫星 l 与 m 之间的最短路径数, $\sigma_{lm}(n)$ 表示经过卫星 n 的最短路径数. 两者归一化后加权融合, 得到节点 n 的拓扑结构因子

$$V_n^{\text{str}}(t) = \omega_1 \cdot \frac{k_n^{\text{shell}}(t)}{\max(k^{\text{shell}})} + \omega_2 \cdot \frac{C_n^{\text{bet}}(t)}{\max(C^{\text{bet}})}, \quad \omega_1 + \omega_2 = 1. \quad (27)$$

(2) 横向传播因子. 该因子用于衡量节点在遭受横向攻击时对全网状态谱半径的影响程度. 根据文献 [34], 卫星网络的渗透传播阈值可表示为 $\rho(\Lambda A - \Omega)$, 其中 Λ 为渗透强度对角矩阵, Ω 为修复强度对角矩阵, A 为网络邻接矩阵. 节点 n 的横向传播因子为

$$V_n^{\text{spr}}(t) = \frac{\partial \rho(\Lambda A - \Omega)}{\partial a_{ij}(t)}, \quad (28)$$

表示节点 n 的局部结构变化对全局传播潜能的边际贡献.

(3) 时序稳定因子. 为刻画节点在动态拓扑下的重要性波动性, 本文引入时序稳定因子

$$V_n^{\text{tmp}}(t) = \exp(-\kappa \cdot \text{Var}(k_n^{\text{shell}}(t))), \quad (29)$$

其中, $\text{Var}(k_n^{\text{shell}}(t))$ 表示节点 n 在给定时间窗口内的度变化方差, κ 为平滑系数. 该指标反映节点在拓扑演化中的稳定性, 度波动越剧烈, 其时序稳定性越低.

综合考虑拓扑结构、横向传播与时序波动三类特征, 卫星 n 的最终重要性得分定义为

$$V_n(t) = \alpha_1 V_n^{\text{str}}(t) + \alpha_2 V_n^{\text{spr}}(t) + \alpha_3 V_n^{\text{tmp}}(t), \quad \alpha_1 + \alpha_2 + \alpha_3 = 1, \quad (30)$$

其中, α_1, α_2 与 α_3 分别为权重系数.

5.3 基于最优控制的最优攻防决策

为求解攻防双方的最优控制策略, 引入协态向量 $\lambda(t) = (\lambda_1(t), \dots, \lambda_N(t))$, 其中 $\lambda_n(t)$ 为卫星 v_n 的协态向量, 定义为 $\lambda_n(t) = [\lambda_n^S(t) \ \lambda_n^L(t) \ \lambda_n^A(t) \ \lambda_n^Q(t) \ \lambda_n^H(t)]^\top$, 构建系统的哈密顿函数如下:

$$\mathcal{H}(X, \lambda, u, \psi, t) = \sum_{n=1}^N \mathcal{L}(x_n, u_n, \psi_n, t) + \lambda_n^\top f(x_n, u_n, \psi_n, t), \quad (31)$$

其中, 瞬时性能指标为

$$\mathcal{L}(x_n, u_n, \psi_n, t) = r_d \psi_n^2 - r_a u_n^2 + c_Q V_n Q_n + c_H V_n H_n - c_L V_n L_n - c_A V_n A_n, \quad (32)$$

状态动力学方程 $f(\cdot)$ 由式 (12)~(16) 定义.

依据 Pontryagin 极大值原理, 攻击者与防御者的控制应分别满足哈密顿函数关于各自控制变量的驻值条件, 即 $\frac{\partial \mathcal{H}}{\partial u_n} = 0$ (对攻击者求极大值) 和 $\frac{\partial \mathcal{H}}{\partial \psi_n} = 0$ (对防御者求极小值). 对 $u_n(t), \forall n \in \mathcal{N}$ 求偏导

$$\frac{\partial \mathcal{H}}{\partial u_n(t)} = -2r_a u_n(t) + \lambda_n^\top \frac{\partial f}{\partial u_n(t)}, \quad (33)$$

$$\frac{\partial f}{\partial u_n(t)} = \left[-\lambda_0^{\text{dir}} S_n(t), \lambda_0^{\text{dir}} S_n(t) - \alpha_0 L_n(t), \alpha_0 L_n(t), 0, 0 \right]^\top. \quad (34)$$

由此可得攻击者最优控制律

$$\begin{aligned} u_n^*(t) &= \Pi_{[0, u_a^{\max}]} \left(\frac{1}{2r_a} \lambda_n^\top \frac{\partial f}{\partial u_n(t)} \right) \\ &= \Pi_{[0, u_a^{\max}]} \frac{1}{2r_a} \{ (\lambda_n^L(t) - \lambda_n^S(t)) \lambda_0^{\text{dir}} S_n(t) + (\lambda_n^A(t) - \lambda_n^L(t)) \alpha_0 L_n(t) \}, \end{aligned} \quad (35)$$

其中, $\Pi_{[0, u_a^{\max}(t)]}(\cdot)$ 为约束区间的映射函数. 对 $\psi_n(t), \forall n \in \mathcal{N}$ 求偏导

$$\frac{\partial \mathcal{H}}{\partial \psi_n(t)} = 2r_d \psi_n(t) + \lambda_n^\top \frac{\partial f}{\partial \psi_n(t)}, \quad (36)$$

$$\frac{\partial f}{\partial \psi_n(t)} = \left[0, -\eta_L L_n(t), -\eta_A A_n(t), \eta_L L_n(t) + \eta_A A_n(t) - \rho_0 Q_n(t), \rho_0 Q_n(t) \right]^\top. \quad (37)$$

由此得到防御者最优控制律

$$\begin{aligned} \psi_n^*(t) &= \Pi_{[0, \psi_d^{\max}]} \left(-\frac{1}{2r_d} \lambda_n^\top \frac{\partial f}{\partial \psi_n(t)} \right) \\ &= \Pi_{[0, \psi_d^{\max}]} \left\{ \frac{1}{2r_d} \left\{ (\lambda_n^L(t) - \lambda_n^Q(t)) \eta_L L_n(t) + (\lambda_n^A(t) - \lambda_n^Q(t)) \eta_A A_n(t) \right. \right. \\ &\quad \left. \left. + (\lambda_n^Q(t) - \lambda_n^H(t)) \rho_0 Q_n(t) \right\} \right\}, \end{aligned} \quad (38)$$

其中, $\Pi_{[0, \psi_d^{\max}(t)]}(\cdot)$ 为约束区间的映射函数. 由二阶导数可知, $\mathcal{H}$ 对 u_n 的二阶偏导为 $-2r_a < 0$, 对 ψ_n 的二阶偏导为 $2r_d > 0$, 因此攻击者的哈密顿函数为凸函数、防御者的为凹函数, 极值点唯一, 最优控制解存在.

根据 Pontryagin 条件, 协态方程可表示为

$$\begin{aligned} \dot{\lambda}_n(t) &= -\frac{\partial \mathcal{H}(x_n, \lambda, u_n, \psi_n, t)}{\partial x_n} = -\frac{\partial \mathcal{L}(x_n, \lambda, u_n, \psi_n, t)}{\partial x_n} - (\partial_{x_n} f)^\top \lambda_n \\ &= \begin{bmatrix} (\lambda_n^S - \lambda_n^L) \left[\lambda_0^{\text{dir}} u_n(t) + \sum_{m=1}^N a_{mn} \lambda_0^{\text{lat}} u_m(t) R_{mn}(t) \right] \\ \alpha_0 (\lambda_n^L - \lambda_n^A) u_n + \eta_L (\lambda_n^L - \lambda_n^Q) \psi_n + c_L V_n \\ \eta_A (\lambda_n^A - \lambda_n^Q) \psi_n + c_A V_n \\ \rho_0 (\lambda_n^Q - \lambda_n^H) \psi_n - c_Q V_n \\ \sigma_H (\lambda_n^H - \lambda_n^S) - c_H V_n \end{bmatrix}. \end{aligned} \quad (39)$$

其终端条件由性能指标的终端项确定

$$\lambda_n(T) = \nabla_{x_n} \Phi(x_n(T)) = \left[-2\gamma_S S_n(T), 2\gamma_L V_n L_n(T), 2\gamma_A V_n A_n(T), -2\gamma_Q V_n Q_n(T), -2\gamma_H V_n H_n(T) \right]^\top. \quad (40)$$

为获得状态轨迹与最优控制决策向量 (u^*, ψ^*) , 基于前述最优控制系统, 本文采用前后向方法 (forward-backward method) [35] 进行迭代求解, 该方法交替更新状态方程与协态方程, 直至收敛至博弈均衡解, 具体见算法 1. 相比于直接求解所构建的动态序贯微分博弈纳什均衡解的指数级复杂度, 算法 1 的整体复杂度为 $\mathcal{O}(MN^2)$, 其中 M 为总时隙数. 由于 LEO 卫星星座在物理拓扑上天然具备层次化分簇结构 [36~38], 其特征表现为卫星在特定轨道面内紧密连接, 而不同轨道面间则通过少量星间链路互联, 因此可将全局防御问题分解为一系列局部簇内的子博弈. 每个簇独立运行算法 1, 仅需计算本簇内节点的防护策略, 并在簇间交换聚合态势指标或边界节点状态信息, 从而将整体复杂度进一步降至 $\mathcal{O}\left(M \sum_{k=1}^K N_k^2\right)$, 其中 K 为簇数, $N_k \ll N$ 为第 k 个簇内节点数.

6 仿真分析

6.1 仿真设置

仿真构建了低轨卫星星座网络拓扑, 分别考虑 10 和 500 颗卫星的两种规模配置. 系统初始状态设为 $x_n(0) = [1, 0, 0, 0, 0]$, $\forall n \in \mathcal{N}$, 表示所有节点初始均处于正常态. 考虑到卫星随轨道运动导致的拓扑动态变化, 采用离散时间方式对卫星位置进行周期性采样, 并在每个采样时刻生成对应的网络拓扑快照. 参考已公开的 Starlink 第一阶

算法 1 三阶段序贯修复博弈求解算法 (three-phase sequential repair game solver).

输入: 时间区间 $[0, T]$, 步长 Δt , 收敛阈值 ε , 松弛因子 $\beta \in (0, 1]$, 最大迭代次数 $K_{\max}$.

- 1: **阶段 I: 网络初始安全态势感知**
- 2: 基于 APT 检测或异常识别方法^[27, 28] 获取系统初始安全态势 $X(0)$;
- 3: 初始化时间网格 $t_0 = 0, t_1, \dots, t_M = T$, 控制轨迹 $u_n^{(0)}(t), \psi_n^{(0)}(t)$, 迭代计数 $k \leftarrow 0$;
- 4: **阶段 II: 动态拓扑下的节点重要性评估**
- 5: **for** 每个时隙 t_k **do**
- 6: 获取邻接矩阵 $A(t_k)$ 与节点状态 $X(t_k)$;
- 7: 基于 k -shell 与介数中心性计算结构感知因子 $V_n^{\text{str}}(t_k)$;
- 8: 基于谱半径敏感度计算传播敏感因子 $V_n^{\text{spr}}(t_k)$;
- 9: 基于度方差平滑计算时序稳定因子 $V_n^{\text{tmp}}(t_k)$;
- 10: 基于式 (30) 综合加权得到节点重要性 $V_n(t_k)$;
- 11: **end for**
- 12: **阶段 III: 基于最优控制与前向 – 后向迭代的最优攻防策略**
- 13: **repeat**
- 14: $k \leftarrow k + 1$;
- 15: 状态正向积分: 使用当前 $u^{(k-1)}(t), \psi^{(k-1)}(t)$ 积分状态方程 (12)~(16);
- 16: 协态初始化: 基于式 (40) 计算 $\lambda_n^{(k)}(T) = \nabla_{x_n} \Phi(x_n^{(k)}(T))$;
- 17: 协态反向积分: 基于式 (39) 从 $t = T$ 到 0 反向积分协态方程 $\dot{\lambda}_n(t) = -\partial \mathcal{H} / \partial x_n$;
- 18: 无约束最优候选控制变量更新:

$$\tilde{u}_n^{(k)}(t) \leftarrow \frac{1}{2r_a} \lambda_n^{(k)\top} \frac{\partial f}{\partial u_n}, \quad (41)$$

$$\tilde{\psi}_n^{(k)}(t) \leftarrow -\frac{1}{2r_d} \lambda_n^{(k)\top} \frac{\partial f}{\partial \psi_n}; \quad (42)$$

- 19: 投影与松弛更新:

$$u_n^{(k)}(t) \leftarrow \Pi_{[0, u_a^{\max}]}((1 - \beta)u_n^{(k-1)}(t) + \beta \tilde{u}_n^{(k)}(t)), \quad (43)$$

$$\psi_n^{(k)}(t) \leftarrow \Pi_{[0, \psi_d^{\max}]}((1 - \beta)\psi_n^{(k-1)}(t) + \beta \tilde{\psi}_n^{(k)}(t)); \quad (44)$$

- 20: **until** 满足收敛条件:

$$\max_{n,t} \{|u_n^{(k)}(t) - u_n^{(k-1)}(t)|, |\psi_n^{(k)}(t) - \psi_n^{(k-1)}(t)|\} < \varepsilon \text{ or } k \geq K_{\max}; \quad (45)$$

输出: 输出最优控制对 $(u^*(t), \psi^*(t))$ 及最优状态轨迹 $X^*(t)$.

段星座设计^[38], 设置轨道高度为 550 km, 轨道倾角为 53° , 星间链路的最大距离基于可见性约束设定为 5014 km, 单星最多可建立 4 条 ISL 链路. 采用随机链路可用性模型, 任意已建立的 ISL 链路 $\{v_m, v_n\}$ 在时刻 t 的连通性以概率 $p_{mn} = 0.5$ 决定. 仿真总步数设置为 1000 步, 在算法 1 中, 最大迭代次数设为 1000, 松弛因子 $\beta = 0.1$, 收敛阈值 $\varepsilon = 1e^{-6}$. 卫星通信模型中参数设置如下^[30]: ISL 链路带宽 $W_{mn} = 80$ MHz, 卫星发射功率 $P_m = 20$ W, 收发天线联合增益 $G_{mn} = 42$ dB, 附加线路损耗 $h_{mn}^a = 2$ dB, 中心频率 $f_{mn}^c = 33.25$ GHz, 噪声温度 $T_0 = 25$ dBK. 为保证评估的公平性, 采用等权重分配策略, 各状态重要性系数统一设为 $\gamma_L = \gamma_A = \gamma_S = \gamma_H = \gamma_Q = 100$, 攻击者与防御者的最大性能约束分别为 $u_d^{\max} = \psi_d^{\max} = 10$, 攻击者成本参数设为 $r_a = c_L = c_A = 1.5$, 防御者成本参数设为 $r_d = c_Q = c_H = 1$, 节点重要性权重系数设为 $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$, 节点结构因子权重取 $\omega_1 = \omega_2 = 1/2$, 其余参数详见表 1. 为评估所提算法的性能, 从防御者的视角设计了以下 3 种对比方案.

(1) 基于 Stackelberg 博弈的 APT 防御策略 (Stackelberg game-based APT defense, SGAD)^[15]. SGAD 策略是一种代表性的基于博弈论的 APT 防御模型, 它将节点状态简化为安全与不安全两类, 攻防双方通过 Stackelberg 博弈框架选择最优策略以最大化各自期望效用.

(2) 固定策略 (fixed strategy). 攻击者采用基于前后向方法的渗透策略, 防御者采用固定防御策略, 设 $\psi_n(t) = 2$. 该基线旨在凸显动态调整防御资源的重要性.

(3) 随机策略 (random strategy). 攻击者采用基于前后向方法的渗透策略, 防御者采用随机防御策略, $\psi_n(t)$

表 1 仿真参数设置.

Table 1 Simulation parameter settings.

Parameter	Description	Value	Parameter	Description	Value
λ_0^{dir}	Direct penetration rate	0.2	λ_0^{lat}	Lateral penetration rate	0.05
α_0	Activation rate	0.5	η_L	Detection efficiency (latent)	0.6
η_A	Detection efficiency (active)	0.7	ρ_0	Recovery rate	0.8
σ_H	Degradation rate	0.1	κ	Temporal stability sensitivity	5

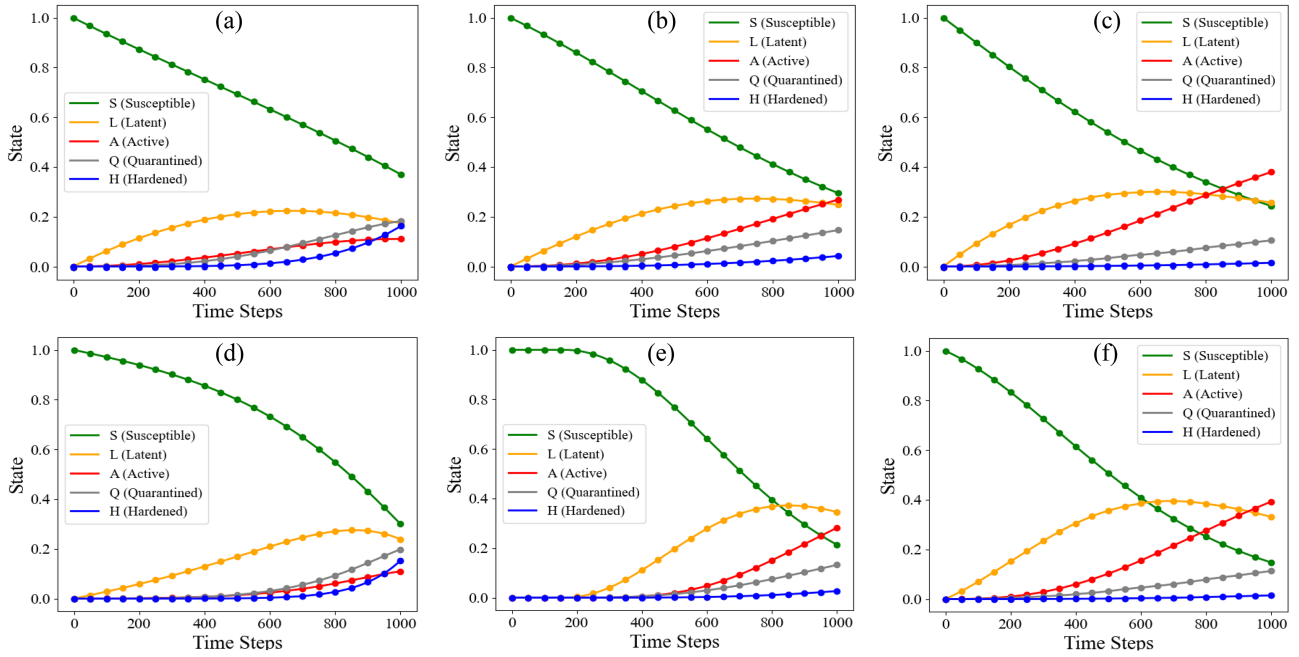


图 3 (网络版彩图) 不同网络规模和防守策略下系统安全状态演变. (a) 所提策略 ($N = 10$); (b) 固定策略 ($N = 10$); (c) 随机策略 ($N = 10$); (d) 所提策略 ($N = 500$); (e) 固定策略 ($N = 500$); (f) 随机策略 ($N = 500$).

Figure 3 (Color online) Security state evolutions under different network scales and defense strategies. (a) The proposed strategy ($N = 10$); (b) the fixed strategy ($N = 10$); (c) the random strategy ($N = 10$); (d) the proposed strategy ($N = 500$); (e) the fixed strategy ($N = 500$); (f) the random strategy ($N = 500$).

在 $[0, 1]$ 内随机取值. 该基线用于验证所提策略相对于盲目防御的效率优势.

仿真中设置了以下 4 类性能评估指标: 系统总效用 (system utility) $U(u, \psi)$ 、安全支配指数 (security dominance index, SDI)、防御掌控指数 (defense control index, DCI) 和平均策略强度 (average strategy intensity). 其中, SDI 用于刻画系统在终端时刻的安全支配程度, 定义为防御者控制状态节点的概率与攻击者控制状态节点的概率之比; DCI 用于衡量防御者在整个博弈过程中对网络节点的总体掌控能力, 当节点处于正常态、隔离态或加固态时视为受防御方掌控, 而处于潜伏态或激活态时视为受攻击方控制; 平均策略强度反映了攻防双方在博弈时间区间 $[0, T]$ 内的资源投入水平, 通过计算控制变量在全时域上的平均值得到.

6.2 仿真结果

为评估所提序贯零和博弈模型的有效性, 首先对比了防御者在不同网络规模和不同策略下系统安全状态演化过程. 如图 3 所示, 当防御者采用所提最优策略时, 能够显著抑制攻击者的渗透与扩散行为, 表现为潜伏态 (L) 与激活态 (A) 节点的比例明显低于采用固定策略或随机策略的方案. 卫星初始状态设置为敏感状态, 随着博弈过程展开, 敏感状态概率逐步降低并向其他状态转移. 如图 3(a) 所示, 隔离态 (Q) 与加固态 (H) 在后期显著上升, 表明防御资源向隔离与加固态倾斜, 从而有效削弱潜伏与激活态节点比例; 相比之下, 图 3(b) 和 (c) 表明, 当

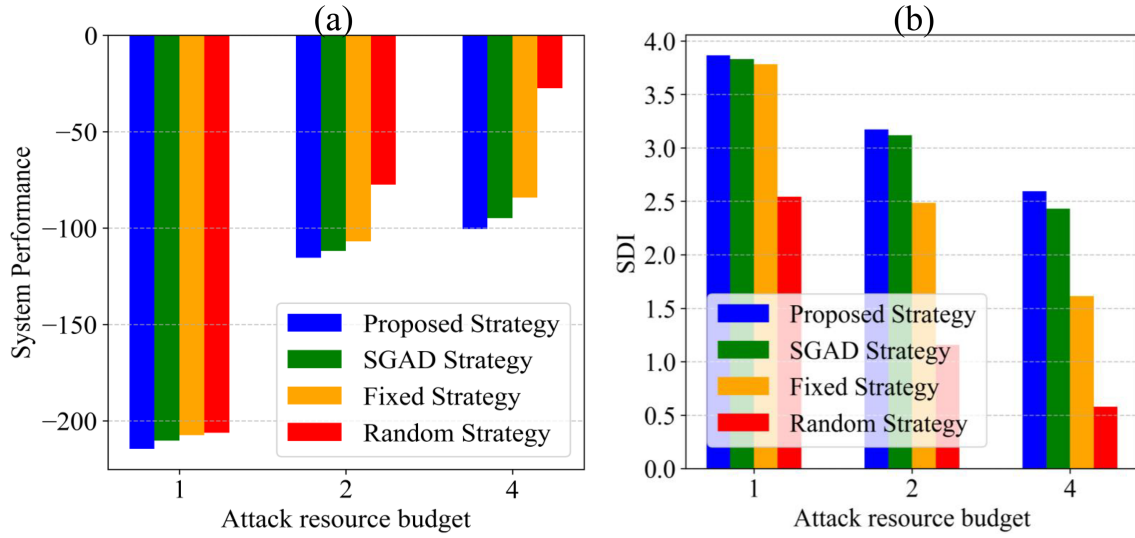


图 4 (网络版彩图) 4 种防守策略下攻击资源预算对系统的影响. (a) 系统效用; (b) 安全支配指数.

Figure 4 (Color online) The impact of attack resource budgets under 4 defense strategies on the system. (a) System utility; (b) security dominance index (SDI).

防御者分别采用固定策略与随机策略时, 难以有效遏制节点向激活态转变的趋势, 后期激活态节点概率仍持续上升, 系统安全性显著下降.

由图 3(d) 可知, 所提动态序贯修复策略在大规模网络中仍能有效抑制攻击传播, 特别地, 潜伏态 (L) 与激活态 (A) 节点比例显著低于对比策略, 且隔离态 (Q) 与加固态 (H) 节点比例稳步上升, 表明防御资源可被自适应地调度至关键节点以阻断 APT 渗透行为. 由图 3(e) 和 (f) 可知, 固定策略与随机策略因难以适应拓扑变化与攻击策略调整, 潜伏态与激活态节点比例随网络规模扩大显著上升, 防御效果随网络规模扩大进一步下降. 与 $N = 10$ 的小规模网络相比, $N = 500$ 的大规模网络中拓扑动态性更为显著, 攻击传播路径更加复杂. 所提策略通过动态节点重要性评估与序贯博弈决策, 仍能保持较低的失陷节点比例, 说明其具备良好的扩展性与适应性.

为进一步验证所提策略的鲁棒性与适应性, 通过限制攻击者的最大策略强度, 构建了不同攻击能力下的对抗场景, 仿真结果如图 4 所示, 在系统效用和安全支配指数方面, 所提策略在所有攻击资源预算下均优于 SGAD 策略及其他基线方案. 由图 4(a) 可见, 当攻击者能力较弱 (即攻击资源预算较低) 时, 各防御策略性能差异不显著; 然而, 随着攻击者能力上限的提升, 系统性能指标整体下降, 而采用所提最优策略的防御者始终保持最优防御效果, 且与其他策略的差距逐渐扩大, 表明所提策略具有较强的适应性与抗攻击能力. 图 4(b) 进一步展示了系统安全性随攻击能力变化的趋势, 即所提防御策略在高强度攻击下仍能维持最高安全性, 而 SGAD 策略、固定与随机策略的安全性下降幅度更大. 由于 SGAD 策略节点状态建模较为简化且仅考虑一轮式两阶段博弈, 未能充分考虑节点状态演化的多维性和不同时间下多轮次修复博弈中攻防策略的动态演变性, 因此其防御策略在应对复杂 APT 攻击时仍显不足; 而所提策略通过精细化的五状态模型和多轮次序贯修复机制, 能更准确地刻画攻防交互过程, 从而实现对系统安全态势的更有效掌控.

为深入分析关键参数对系统性能、安全和最优决策的影响, 进一步开展了 3 组敏感性分析实验.

直接渗透率与横向传播率的影响. 由图 5 可知, 当直接渗透率增大时, 攻击者通过直接渗透攻陷节点的能力增强, 系统性能指标呈现显著下降趋势. 相应地, 横向传播率的增大意味着攻击在卫星网络内部的扩散能力变强, 导致系统性能指标下降.

防御者对终端时刻状态重视程度的影响. 图 6(a) 和 (b) 通过调整与安全状态相关的终端时刻状态重要性系数 $\gamma_S, \gamma_H, \gamma_Q$, 其取值范围为 100~1000, 步长为 50, 用以分析其对防御掌控指数和系统效用的影响. 由图 6(a) 可见, 随着终端时刻状态重视程度的提升, 所提最优策略能够自适应地增加防御资源投入, 从而显著提升系统防御掌控能力. 相较之下, 固定与随机策略由于缺乏动态调整能力, 当攻击者强化对终端时刻目标的攻击强度时,

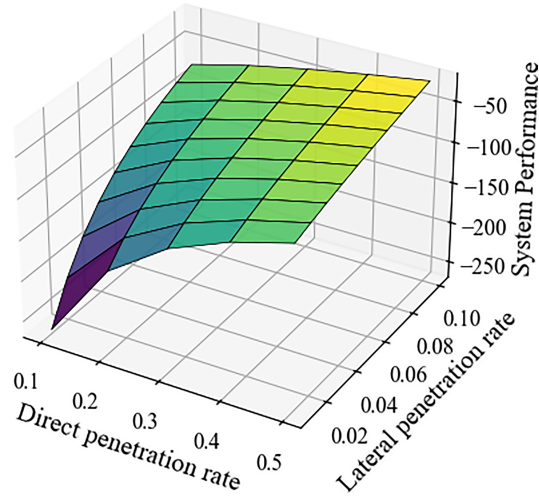


图 5 (网络版彩图) 直接渗透率与横向传播率对系统性能的影响。

Figure 5 (Color online) The impact of direct & lateral penetration rates on the system.

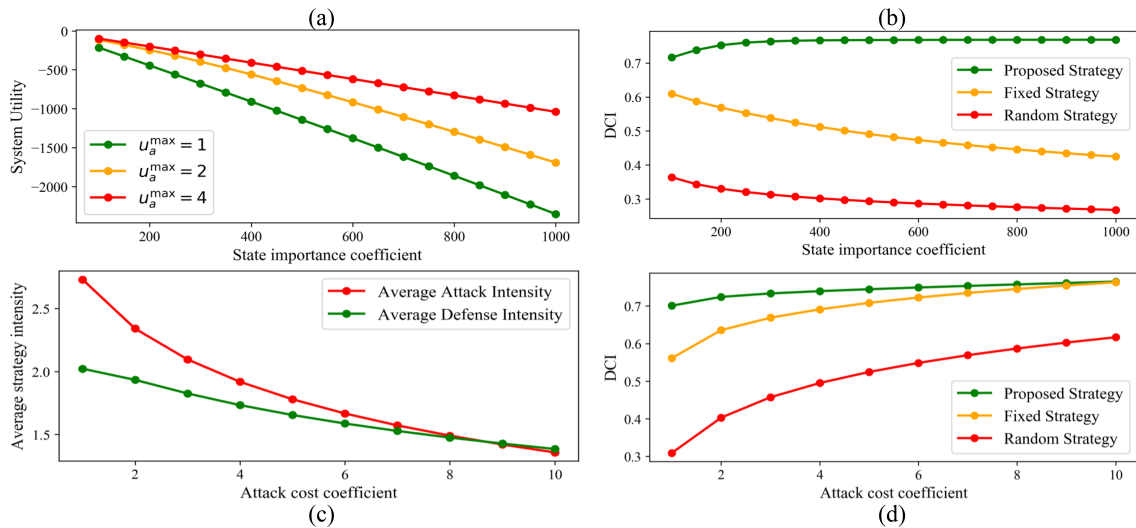


图 6 (网络版彩图) 状态重要性系数和攻击成本系数对系统的影响。(a) 系统效用与状态重要性系数;(b) 防御掌控指数与状态重要性系数;(c) 平均攻防强度与攻击成本系数;(d) 防御掌控指数与攻击成本系数。

Figure 6 (Color online) The impact of state importance coefficient and attack cost coefficient on the system. (a) System utility vs. state importance coefficient; (b) defense control index (DCI) vs. state importance coefficient; (c) average attack/defense intensity vs. attack cost coefficient; (d) DCI vs. attack cost coefficient.

防御掌控指数反而下降。图 6(b) 进一步展示了终端时刻状态重视程度对系统性能的影响。随着防御者对终端时刻状态重视程度的增加, 攻击者整体收益持续下降, 原因在于防御收益系数的增强; 在不同场景对比中, 对弱攻击者的防御收益提升更为显著。

攻击者成本系数 r_a 的影响。 图 6(c) 和 (d) 分析了攻击成本系数 r_a 在区间 $[1, 10]$ 内 (步长为 1) 变化时对系统攻防策略的影响。由图 6(c) 可见, 随着攻击成本的上升, 攻击方的平均渗透策略强度逐渐减弱, 防御者的最优修复策略强度亦呈相应下降趋势, 表明所提最优防御策略能够自适应地响应攻击者策略变化。图 6(d) 展示了攻击成本对系统掌控度的影响: 攻击成本的增加有效遏制了攻击行为, 使得终端时刻防御掌控指数在 3 种策略下均得到提升。值得注意的是, 尽管固定与随机策略在此情形下也表现出较高的防御掌控指数, 但这主要源于其在整个过程中维持高水平、静态的防御投入, 导致其整体防御资源利用效率偏低。

7 结论

本文针对卫星互联网在面临 APT 威胁时的动态安全防护问题,提出了一种基于攻防博弈的时变卫星网络最优修复策略.考虑卫星网络的网络拓扑动态、安全状态演变与防御资源受限特性,构建了节点安全状态演化模型和序贯修复博弈模型,来刻画 APT 攻击的渐进渗透与防御者的动态修复决策,提出了动态拓扑下的卫星节点重要性评估算法,并结合最优控制理论和前后向方法设计了分层快速近优均衡求解算法,实现了低开销自适应的最优防御策略决策.仿真结果表明,该方法能在多种攻防配置场景下显著降低 APT 攻击造成的系统损害,提升网络恢复效率与防御鲁棒性.后续研究将进一步探索多维信息不对称以及有限理性下的跨星座动态安全博弈与防护策略.

参考文献

- Zhang G X, Wang Y F, Ding X J, et al. Research on several key technologies of satellite Internet. *J Commun*, 2021, 42: 1–14 [张更新, 王运峰, 丁晓进, 等. 卫星互联网若干关键技术研究. *通信学报*, 2021, 42: 1–14]
- Ni S J, Yue Y, Zuo Y, et al. The status quo and prospect of satellite network routing technology. *J Electron Inform Technol*, 2023, 45: 383–395 [倪少杰, 岳洋, 左勇, 等. 卫星网络路由技术现状及展望. *电子与信息学报*, 2023, 45: 383–395]
- Yue P, An J, Zhang J, et al. Low earth orbit satellite security and reliability: issues, solutions, and the road ahead. *IEEE Commun Surv Tut*, 2023, 25: 1604–1652
- Viasat. Case study. 2022. <https://cyberconflicts.cyberpeaceinstitute.org/law-and-policy/cases/viasat>
- Zhang L, Zhu T, Hussain F K, et al. A game-theoretic method for defending against advanced persistent threats in cyber systems. *IEEE Trans Inform Forensic Secur*, 2022, 18: 1349–1364
- Li H, Wu J, Xu H, et al. Explainable intelligence-driven defense mechanism against advanced persistent threats: A joint edge game and AI approach. *IEEE Trans Depend Secure Comput*, 2022, 19: 757–775
- He W, Tan J, Guo Y, et al. A deep reinforcement learning-based deception asset selection algorithm in differential games. *IEEE Trans Inform Foren Secur*, 2024, 19: 8353–8368
- Zoukarni A, Damara S S, Praveen Kumar M S, et al. Defending against APT attacks in robots: A multi-phase game-theoretical approach. In: *Proceedings of Decision and Game Theory for Security*, 2025. 287–305
- Moothedath S, Sahabandu D, Allen J, et al. Dynamic information flow tracking for detection of advanced persistent threats: a stochastic game approach. *IEEE Trans Automat Contr*, 2024, 69: 6684–6699
- Shen S, Cai C, Shen Y, et al. Joint mean-field game and multiagent asynchronous advantage actor-critic for edge intelligence-based IoT malware propagation defense. *IEEE Trans Depend Secure Comput*, 2025, 22: 3824–3838
- Hu Y, Chen J, Zhu Q. Game-theoretic neyman-pearson detection to combat strategic evasion. *IEEE Trans Inform Foren Secur*, 2024, 20: 516–530
- Wang Z, Duan C J, Wu T, et al. Research on optimizing security control mechanism of networked system based on Stackelberg defender-attacker game. *J Cyber Secur*, 2019, 4: 101–115 [王震, 段晨健, 吴铤, 等. 基于 Stackelberg 攻防博弈的网络系统安全控制机制优化研究. *信息安全学报*, 2019, 4: 101–115]
- Shen S, Xie L, Zhang Y, et al. Joint differential game and double deep Q-networks for suppressing malware spread in Industrial Internet of Things. *IEEE Trans Inform Forensic Secur*, 2023, 18: 5302–5315
- Yang L X, Li P, Zhang Y, et al. Effective repair strategy against advanced persistent threat: a differential game approach. *IEEE Trans Inform Forensic Secur*, 2018, 14: 1713–1728
- Bi J, He S, Luo F, et al. Defense of advanced persistent threat on Industrial Internet of Things with lateral movement modeling. *IEEE Trans Ind Inf*, 2023, 19: 9619–9630
- Han C, Huo L, Tong X, et al. Spatial anti-jamming scheme for internet of satellites based on the deep reinforcement learning and stackelberg game. *IEEE Trans Veh Technol*, 2020, 69: 5331–5342
- Han C, Liu A, Wang H, et al. Dynamic anti-jamming coalition for satellite-enabled army IoT: a distributed game approach. *IEEE Int Things J*, 2020, 7: 10932–10944
- Zhang X, Zhang B, Guo D, et al. Potential game-based radio resource allocation in uplink multibeam satellite IoT networks. *IEEE Trans Aersp Electron Syst*, 2021, 57: 4269–4279
- Liao C, Xu K, Zhu H, et al. Secure transmission in satellite-UAV integrated system against eavesdropping and jamming: a two-level stackelberg game model. *China Commun*, 2022, 19: 53–66
- Cui Y, Xi Z, Zhang X. A game between cyber attack and defense under the topology of satellite networks. In: *Proceedings of Advances in Guidance, Navigation and Control*, 2023

- 21 Vasquez C. How a hacking crew overtook a satellite from inside a Las Vegas convention center and won \$50,000. 2023. <https://cyberscoop.com/mhackeroni-hackasat-space-def-con/>
- 22 Yin J, Li Y H, Liao S K, et al. Entanglement-based secure quantum cryptography over 1,120 kilometres. *Nature*, 2020, 582: 501–505
- 23 Li H, Shi D, Wang W, et al. Secure routing for LEO satellite network survivability. *Comput Netws*, 2022, 211: 109011
- 24 Al Homssi B, Dakic K, Wang K, et al. Artificial intelligence techniques for next-generation massive satellite networks. *IEEE Commun Mag*, 2023, 62: 66–72
- 25 Huang C, Wang D X, Hou R H. Satellite security transmission optimization algorithm for joint anomaly detection. *Acta Electron Sin*, 2025, 53: 1460–1468 [黄晨, 王定轩, 侯蓉晖. 联合异常检测的卫星安全传输优化算法. *电子学报*, 2025, 53: 1460–1468]
- 26 Wu H, Li J, Lu H, et al. A two-layer caching model for content delivery services in satellite-terrestrial networks. In: *Proceedings of IEEE Global Communications Conference (GLOBECOM)*, 2016. 1–6
- 27 Milajerdi S M, Gjomemo R, Eshete B, et al. HOLMES: real-time APT detection through correlation of suspicious information flows. In: *Proceedings of IEEE Symposium on Security and Privacy (SP)*, 2019. 1137–1152
- 28 Xiong C, Zhu T, Dong W, et al. Conan: a practical real-time APT detection system with high accuracy and efficiency. *IEEE Trans Dependable Secure Comput*, 2022, 19: 551–565
- 29 Xu G, Tan F, Ran Y, et al. Joint beam-hopping scheduling and coverage control in multibeam satellite systems. *IEEE Wireless Commun Lett*, 2023, 12: 267–271
- 30 Fan K, Feng B, Zhang X, et al. Demand-driven task scheduling and resource allocation in space-air-ground integrated network: a deep reinforcement learning approach. *IEEE Trans Wireless Commun*, 2024, 23: 13053–13067
- 31 Xu Z, Chen J, Wang B C, et al. Data-driven mean field equilibrium computation in large-population LQG games. *IEEE Trans Control Netw Syst*, 2025, 12: 2713–2725
- 32 Wang Y, Su Z, Benslimane A, et al. Collaborative honeypot defense in UAV networks: a learning-based game approach. *IEEE Trans Inform Forensic Secur*, 2024, 19: 1963–1978
- 33 Maji G, Namtirtha A, Dutta A, et al. Influential spreaders identification in complex networks with improved k-shell hybrid method. *Expert Syst Appl*, 2020, 144: 113092
- 34 Zino L, Cao M. Analysis, prediction, and control of epidemics: a survey from scalar to dynamic network models. *IEEE Circ Syst Mag*, 2021, 21: 4–23
- 35 Shampine L F. *Numerical Solution of Ordinary Differential Equations*. New York: Routledge, 2018
- 36 Jiao J, Yang P, Du Z, et al. Clustered multi-criteria routing algorithm for mega low earth orbit satellite constellations. *IEEE Trans Veh Technol*, 2024, 73: 13790–13803
- 37 Zhang C, Yin T, Wang W, et al. Clustering strategy for megaconstellation with synergistic energy and size considerations. *IEEE Trans Aerosp Electron Syst*, 2025, 61: 10098–10109
- 38 Bhattacharjee D, Singla A. Network topology design at 27,000 km/hour. In: *Proceedings of the International Conference on Emerging Networking Experiments And Technologies (CoNEXT)*, 2019. 341–354

Optimal recovery strategies for satellite Internet under advanced persistent threats

Yuntao WANG, Chunlin HE & Zhou SU*

School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China

* Corresponding author. E-mail: zhousu@xjtu.edu.cn

Abstract With the rapid development of low Earth orbit (LEO) satellite constellations and space information infrastructures, the satellite Internet has become a crucial component of the global information network. However, due to its time-varying links, heterogeneous satellite-ground architecture, and resource-constrained nodes, it exhibits high stealth and defense complexity under advanced persistent threats (APTs). Traditional game-theoretic defense strategies, typically designed for static or semi-static topologies, are inadequate in such highly dynamic environments. To address these challenges, this paper proposes a dynamic and optimal recovery strategy tailored for time-varying satellite networks. First, a dynamic sequential repair game model is established to characterize the adversarial process between the attacker and defender, in which the multi-stage evolution of APT penetration and dynamic defense actions is modeled as a series of sequential zero-sum subgames over a finite time horizon. Then, considering the dynamic evolution of network topology, node heterogeneity, and limited defense resources, an optimal control-based forward-backward search and strategy update algorithm is developed to efficiently compute the game equilibrium with low computational complexity. Simulation results demonstrate that the proposed method effectively mitigates cumulative attack damage and enhances defense robustness under resource constraints.

Keywords satellite Internet security, advanced persistent threats, dynamic game, optimal recovery