doi:10.3772/j.issn.1002-0470.2024.02.006

非完全信息下协作式入侵检测系统检测库配置研究①

石月楼② 杨旦杰 冯 宇 李永强

(浙江工业大学信息工程学院 杭州 310023)

摘 要 本文研究了有限时间下非完全信息协作式入侵检测系统(IDS)的检测库配置。针对各入侵检测系统面对不同类型攻击时的最优检测库的配置以及检测库分配的矛盾,提出了一种双层检测库配置方法。第1层研究的是各攻击者的策略制定以及对应入侵检测系统的最优检测库配置方案;第2层研究的是如何解决多个入侵检测系统加载同一检测库的矛盾,给出了基于策略共享的集中式分配方法。第1层研究又可分为2步解决:第1步针对攻击者无法获知系统的真实状态、与入侵检测系统之间存在着信息不对称的问题,提出构建一个基于信念的随机博弈框架,通过后向递归算法求解,并证明该解就是博弈的稳态纳什均衡(SNE)策略;第2步通过求解混合 Markov 决策过程得到各入侵检测系统的最优检测库配置方案。仿真结果表明,本文所提方法能有效地得到协作式入侵检测系统面对不同类型攻击时的最优检测库配置方案。

关键词 非完全信息博弈;Markov 决策过程;入侵检测系统(IDS);资源分配;网络安全

近年来,随着计算机网络和信息技术的不断发展,“黑客攻击”已成为威胁互联网安全的主要因素之一,如何有效地防御恶意入侵、保证网络安全受到了广泛的研究[1-3]。

入侵检测系统(intrusion detection system, IDS)是一种动态检测网络的安全防御机制,可以通过运行入侵检测算法来识别网络中的攻击,有效弥补安全防护技术中的不足[4]。常用的入侵检测算法有异常检测算法[5]、误用检测算法[6]和人工智能算法[7]。然而单一算法无法捕获所有的攻击,且由于能量的限制,单个入侵检测系统无法同时运行所有的算法,因此大量的研究工作投入到了入侵检测系统的检测库配置问题上。其中,博弈论[8-9]是分析该问题的有效工具。文献[10]将入侵检测系统与攻击者的交互过程建模为一个两人非零和随机博弈,通过求解效用最优化方程得到攻防双方的最优策略。文献[11]也用相同的博弈框架研究了类似的问题,并通过强化学习算法求得攻击者和入侵检测系统的最优策略。文献[12]通过两人零和随机博弈框架研究了入侵检测系统检测库配置的问题,提出一种基于强化学习的算法来得到最优检测库配置策略。在文献[13]中,作者使用零和贝叶斯信号博弈研究了入侵检测系统与未知类型的攻击者的交互过程,并证明博弈的纳什均衡策略就是一组最优攻防策略。为了可以运行更多的算法提升检测性能,越来越多的研究提出在多个入侵检测系统中引入协作机制。文献[14]验证了协作式入侵检测系统可以显著提升整体的检测性能,并提出分布式激励分配方法来解决检测库分配的矛盾。文献[15]也研究了类似的问题,并且还考虑了传输延迟和资源有限对检测库分配的影响。文献[16]在文献[15]的基础上还考虑了攻击者之间的协作性,提出了一种基于演化博弈的入侵检测系统检测资源分配模型。求解博弈均衡的方法有 Sarsa 算法[17]、Q-learning 算

① 国家自然科学基金(61973276)资助项目。
② 女,1996 年生,硕士生;研究方向:网络安全;联系人,E-mail: 2111903038@zjut.edu.cn。
(收稿日期:2022-12-18)

法[18]、Monte Carlo 算法[19]等,这些都属于强化学习范畴[20]。强化学习的应用场景非常广泛,尤其适用于环境动力学无法建模的场景,智能体通过与环境的交互,将得到的奖励值作为反馈信号进行训练[21],且泛化性好,不需要进行大量调参。但传统的强化学习算法只能处理离散状态,而本文中的信念状态是连续的,因此本文提出后向递归算法来求解博弈问题。

与现有的大多数研究完全信息下单个入侵检测系统配置不同,本文考虑了一个非完全信息下协作式入侵检测系统检测库配置的问题。入侵检测系统可以监测到系统的状态,而攻击者无法获知,只能通过信念估计,因此通过非完全信息博弈模拟各入侵检测系统和对应攻击者的交互过程符合该情况。并且由于多个入侵检测系统共用一套检测库,因此在得到最优检测库配置方案后还要考虑同一检测库被重复加载的矛盾。本文主要工作总结如下:首先,为了解决信息不对称的情况,攻击者通过假想一个入侵检测系统建立基于信念的随机博弈框架来制定策略,通过后向递归算法求解,并证明了有限时间非完全信息博弈中稳态纳什均衡(stationary Nash equilibrium, SNE)的存在;然后,由于入侵检测系统可以观察到攻击者的策略,因此最优检测库配置方案可以通过求解一个混合 Markov 决策过程得到;最后,提出了一种基于策略共享的集中式分配方法解决检测库被重复加载的矛盾。

1 问题定义

如图 1 所示,入侵检测网络(intrusion detection network,IDN)由 N 个相互连接的入侵检测系统 IDS_i ($i \in \{1,2,\cdots,N\}$) 组成,对应的攻击者用 γ_i ($i \in \{1,2,\cdots,N\}$) 表示。每个子系统的状态集合都是 $S = \{S_1,\cdots,S_q,\cdots,S_K\}$,其中,$S_q$ 表示某种特定运行状态;例如,S_1 表示运行正常,S_2 表示正遭受恶意攻击,S_3 表示运行异常。每个攻击者的纯动作集合都是 $A_\gamma = \{1,\cdots,h,\cdots,H\}$,攻击者 γ_i 发动 h 类型的攻击的能耗为 $c_{\gamma_i}(h) > 0$。

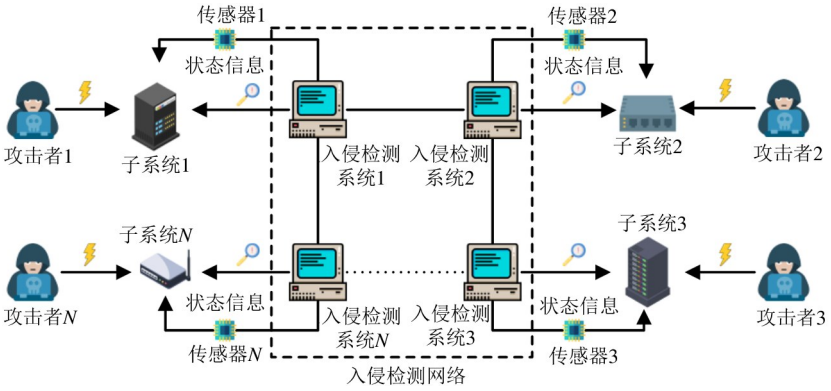

图 1 协作式入侵检测系统模型图

协作式入侵检测系统共用的检测库集合为 $L = \{l_1,\cdots,l_m,\cdots,l_H\}$,检测库 l_m 成功检测到 h 类型的攻击的概率是 $p_{l_m,h}$。当加载的检测库与攻击类型相匹配时,成功检测到的概率会大幅提升。入侵检测系统可以通过加载多个检测库来优化检测性能,但是相应的检测能耗也会变高,因此要考虑检测准确率和能耗之间的一个平衡。每个入侵检测系统每次可以同时加载 n ($n \in \{1,2,\cdots,H\}$) 个检测库,则检测库 L 的可能组合类型共有 $M = C_H^n$ 种,C_H^n 表示从 H 个检测库中挑出 n 个检测库的所有可能组合数,用 F_j ($j \in \{1,2,\cdots,M\}$) 表示,因此,各入侵检测系统的纯动作(检测库配置方案)集合都可以用 $A_{\mathrm{IDS}} = \{F_1,F_2,\cdots,F_M\}$ 表示。令入侵检测系统 IDS_i 加载检测库 l_m 的能耗为 $\bar{c}_{\mathrm{IDS}_i}(l_m) > 0$,则入侵检测系统 IDS_i 选择配置方案 $a_{\mathrm{IDS}_i} \in A_{\mathrm{IDS}}$ 的总能耗是 $c_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i}) = \sum_{l_m \in a_{\mathrm{IDS}_i}} \bar{c}_{\mathrm{IDS}_i}(l_m)$。

接下来,定义入侵检测系统 IDS_i 选择动作 a_{IDS_i} 成功检测到攻击 a_{γ_i} 的概率为

$$p_{a_{\mathrm{IDS}_i},\,a_{\gamma_i}} = 1 - \prod_{l_m \in a_{\mathrm{IDS}_i}} (1 - p_{l_m,\,a_{\gamma_i}}) \tag{1}$$

当子系统 i 状态为 S_k 时，如果攻击 a_{γ_i} 没有被成功检测到，其损失为 $D(S_k,\ a_{\gamma_i})$，则部署在该子系统上的入侵检测系统 IDS_i 的损失为

$$\begin{aligned}\omega(S_k, a_{\mathrm{IDS}_i},\ a_{\gamma_i}) &= (1 - p_{a_{\mathrm{IDS}_i},a_{\gamma_i}})D(S_k,\ a_{\gamma_i}) \\ &\quad + \delta_{\mathrm{IDS}_i} c_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i}) - \delta_{\gamma_i} c_{\gamma_i}(a_{\gamma_i})\end{aligned} \tag{2}$$

其中，$\delta_{\mathrm{IDS}_i} > 0$ 表示与 IDS_i 检测能耗相关的权重参数，$\delta_{\gamma_i} > 0$ 表示与攻击能耗相关的权重参数。最后，定义每个子系统状态转移矩阵为

$$\boldsymbol{T}(a_{\mathrm{IDS}_i}, a_{\gamma_i}) = [\boldsymbol{T}_{S_e,S_f}(a_{\mathrm{IDS}_i},\ a_{\gamma_i})]_{e,f=1,1}^{K,K} \tag{3}$$

其中，$\boldsymbol{T}_{S_e,S_f}(a_{\mathrm{IDS}_i},\ a_{\gamma_i})$ $(e,f \in \{1,2,\cdots,K\})$ 表示子系统 i 在联合纯动作 $\{a_{\mathrm{IDS}_i},\ a_{\gamma_i}\}$ 下从当前状态 S_e 转移到下一状态 S_f 的概率。

由于协作式入侵检测系统共用一套检测库，因此，任一检测库在每一时刻只能被分配给一个入侵检测系统，即各入侵检测系统的检测库配置方案 $\{a_{\mathrm{IDS}_1}, a_{\mathrm{IDS}_2}, \cdots, a_{\mathrm{IDS}_N}\}$ 需满足：

$$a_{\mathrm{IDS}_1} \cap a_{\mathrm{IDS}_2} \cap \cdots \cap a_{\mathrm{IDS}_N} = \varnothing \tag{4}$$

$$a_{\mathrm{IDS}_1} \cup a_{\mathrm{IDS}_2} \cup \cdots \cup a_{\mathrm{IDS}_N} \subseteq L \tag{5}$$

2 非完全信息入侵检测系统检测库配置

本文聚焦非完全信息下协作式入侵检测系统的检测库配置，接下来将分 3 步研究这个问题。首先，构建一个基于信念的随机博弈来研究攻击者的策略制定问题；然后，通过混合 Markov 决策过程得到入侵检测系统的检测库配置策略；最后，基于策略共享的检测库分配方法来解决加载检测库的矛盾。

2.1 攻击策略

每个入侵检测系统和对应攻击者的交互过程都可以描述为一个两人零和随机博弈，动作集、回报函数、子系统的状态集合以及状态转移概率都是相同的，因此在以下建立的基于信念的随机博弈框架中，只考虑其中一个入侵检测系统和其对应攻击者之间的交互，其余 $N-1$ 对的交互过程也是完全一样的。攻击者 γ_i 无法获得子系统的真实状态，通过信念估计，并假想一个同样使用信念的入侵检测系统 $\overline{IDS_i}$ 来制定策略，建立一个由五元组 $<I,A,B,T,R>$ 组成的基于信念的随机博弈框架。

(1)参与者：令 $I = \{\overline{IDS_i}, \gamma_i\}$ $(i \in \{1,2,\cdots,N\})$ 为博弈的参与者集合，其中，$\overline{IDS_i}$ 表示假想的入侵检测系统，γ_i 表示对应的攻击者。

(2)动作：令 $A = \Delta(A_{\mathrm{IDS}}) \times \Delta(A_\gamma)$ 表示纯动作集合 $A_{\mathrm{IDS}} \times A_\gamma$ 上的联合概率分布集合。

(3)信念：令 $B = \Delta(S)$ 表示信念集合，$\Delta(S)$ 表示在状态集 S 上的概率分布。记 $B_t(k)$ 表示 t 时刻子系统 i 状态为 S_k $(k \in \{1,2,\cdots,K\})$ 的概率。在 t 时刻末，攻击者 γ_i 根据观察到的动作 $\{a^o_{\mathrm{IDS}_i,t},\ a^o_{\gamma_i,t}\}$ 用下式更新 $t+1$ 时刻的信念：

$$B_{t+1} = B_t \boldsymbol{T}(a^o_{\mathrm{IDS}_i,t},\ a^o_{\gamma_i,t}) \tag{6}$$

其中，$\boldsymbol{T}(a^o_{\mathrm{IDS}_i,t},\ a^o_{\gamma_i,t})$ 是在式(3)定义的状态转移矩阵。

(4)信念转移概率：令 t 时刻信念为 $b \in B$，联合概率动作为 $\beta = \{\beta_{\overline{\mathrm{IDS}_i}}, \beta_{\gamma_i}\} \in A$，则 $t+1$ 时刻信念转移到 $b' \in B$ 的概率可以表示为 $T(b' \mid b,\beta) = \Pr(B_{t+1} = b' \mid B_t = b,\ A_t = \beta)$。根据信念更新式(6)，当 $b' = b\boldsymbol{T}(a_{\overline{\mathrm{IDS}_i}},\ a_{\gamma_i})$ 时，有 $T(b' \mid b,\beta) = \Pr(a^o_t = \bar{a} \mid B_t = b,\ A_t = \beta)$；否则，$T(b' \mid b,\beta) = 0$。于是可以得到：

$$\Pr(a^o_t = \bar{a} \mid B_t = b,\ A_t = \beta) = \beta_{\overline{\mathrm{IDS}_i}}(a_{\overline{\mathrm{IDS}_i}})\beta_{\gamma_i}(a_{\gamma_i}) \tag{7}$$

其中，$\bar{a} = \{a_{\overline{\mathrm{IDS}_i}},\ a_{\gamma_i}\}$ 表示 t 时刻末观察到的加权入侵检测系统和攻击者的动作。

(5)收益：令 $r_{\gamma_i}(B_t = b, A_t = \beta)$ 表示 t 时刻攻击者 γ_i 的收益，是根据信念 b 对所有状态下收益的加权和，具体表达式为

$$\begin{aligned}r_{\gamma_i}(B_t = b, A_t = \beta) &= \sum_{a_{\overline{\mathrm{IDS}_i}} \in A_{\mathrm{IDS}}} \sum_{a_{\gamma_i} \in A_\gamma} \beta_{\overline{\mathrm{IDS}_i}}(a_{\overline{\mathrm{IDS}_i}}) \\ &\quad \times \beta_{\gamma_i}(a_{\gamma_i}) \sum_{j=1}^{K} b(j)\omega(j,\ a_{\overline{\mathrm{IDS}_i}},\ a_{\gamma_i})\end{aligned} \tag{8}$$

其中，$\omega(j, a_{\overline{\mathrm{IDS}_i}},\ a_{\gamma_i})$ 已在式(2)定义。由零和博弈的性质可知攻击者 γ_i 的收益是加权入侵检测系统 $\overline{\mathrm{IDS}_i}$ 的损失，因此加权入侵检测系统在 t 时刻的收益函数可以表示为 $r_{\overline{\mathrm{IDS}_i}}(B_t = b,\ A_t = \beta) = -r_{\gamma_i}(B_t = b, A_t = \beta)$。又因为收益函数是时不变的，因此可

以简写为 $r_{\gamma_i}(b,\beta)$ 和 $r_{\overline{\mathrm{IDS}_i}}(b,\beta)$。

本文研究的是稳态策略,记为从信念空间 B 到动作空间 A 的映射,即 $\pi^N: B \to A$。在给定信念 $b \in B$, 稳态策略 $\pi^N(b) = \{\pi^N_{\overline{\mathrm{IDS}_i}}(b), \pi^N_{\gamma_i}(b)\}$ 是定义在纯动作集合 $A_{\mathrm{IDS}} \times A_\gamma$ 上的一个概率分布。因此,加权入侵检测系统 $\overline{IDS_i}$ 和对应攻击者 γ_i 的累计收益函数分别为

$$J^N_{\overline{\mathrm{IDS}_i}}(b_0, \pi^N(b)) = \sum_{t=0}^{N} r_{\overline{\mathrm{IDS}_i}}(b, \pi^N(b)) \tag{9}$$

$$J^N_{\gamma_i}(b_0, \pi^N(b)) = \sum_{t=0}^{N} r_{\gamma_i}(b, \pi^N(b)) \tag{10}$$

其中, b_0 是初始信念。下面给出有限时间非完全信息博弈问题的定义。

问题 1 令 $G_N := (I, A, B, T, R)$ 表示一个有限时间非完全信息博弈,求解该博弈问题就等价于求解稳态纳什均衡策略 $\pi^{N*}(b) := \{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b), \pi^{N*}_{\gamma_i}(b)\}$ 满足:

$$J^N_{\overline{\mathrm{IDS}_i}}(b_0, \{\pi^N_{\overline{\mathrm{IDS}_i}}(b), \pi^{N*}_{\gamma_i}(b)\}) \leqslant J^N_{\overline{\mathrm{IDS}_i}}(b_0, \{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b), \pi^{N*}_{\gamma_i}(b)\}) \tag{11}$$

$$J^N_{\gamma_i}(b_0, \{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b), \pi^N_{\gamma_i}(b)\}) \leqslant J^N_{\gamma_i}(b_0, \{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b), \pi^{N*}_{\gamma_i}(b)\}) \tag{12}$$

2.2 入侵检测系统的检测库配置方案

由于真正的入侵检测系统在信息方面具有优势,因此在博弈的每个阶段,攻击者的策略制定过程对入侵检测系统而言是公开的,这使得入侵检测系统能够根据攻击者的策略来制定自己的策略。将该过程建立为一个混合 Markov 决策过程,并用一个四元组 $< \hat{A}, U, \hat{T}, r_{\mathrm{IDS}_i} >$ 来描述。

(1)动作:令 $\hat{A} = \Delta(A_{\mathrm{IDS}}) \times \Delta(A_\gamma)$ 是在纯动作集合 $A_{\mathrm{IDS}} \times A_\gamma$ 上的联合概率分布集合。

(2)状态:令 $U = S \times B$ 为混合状态空间,其中, S 是子系统的状态空间, B 是在基于信念的随机博弈中定义的信念空间。

(3)混合状态转移概率:令 t 时刻混合状态为 $u \in U$, 联合概率动作为 $\hat{\beta} = \{\beta_{\mathrm{IDS}_i}, \beta_{\gamma_i}\} \in \hat{A}$, 则 $t+1$ 时刻混合状态转移到 $u' \in U$ 的概率为 $\hat{T}(u' \mid u, \hat{\beta}) = \Pr(U_{t+1} = u' \mid U_t = u, \hat{A}_t = \hat{\beta})$。

(4)收益:入侵检测系统 IDS_i 的收益函数为

$$r_{\mathrm{IDS}_i}(u, \hat{\beta}) = \sum_{a_{\mathrm{IDS}_i} \in A_{\mathrm{IDS}}} \sum_{a_{\gamma_i} \in A_\gamma} \beta_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i}) \beta_{\gamma_i}(a_{\gamma_i}) \times \omega(s, a_{\mathrm{IDS}_i}, a_{\gamma_i}) \tag{13}$$

其中, $\omega(s, a_{\mathrm{IDS}_i}, a_{\gamma_i})$ 已在式(2)定义。

入侵检测系统 IDS_i 的稳态策略用 $\tau^N_{\mathrm{IDS}_i}$ 表示,是混合状态空间 U 到动作空间 $\Delta(A_{\mathrm{IDS}})$ 的映射,即 $\tau^N_{\mathrm{IDS}_i}: U \to \Delta(A_{\mathrm{IDS}})$。入侵检测系统 IDS_i 的累计收益函数为

$$J^N_{\mathrm{IDS}_i}(u_0, \tau^N_{\mathrm{IDS}_i}(u)) = \sum_{t=0}^{N} r_{\mathrm{IDS}_i}(u, \{\tau^N_{\mathrm{IDS}_i}(u), \pi^N_{\gamma_i}(b)\}) \tag{14}$$

其中, u_0 是初始混合状态, $\pi^N_{\gamma_i}(b)$ 是攻击者 γ_i 的稳态策略。在获知攻击者 γ_i 的稳态纳什均衡策略之后,入侵检测系统 IDS_i 的目标是最大化累计收益函数,因此,入侵检测系统的最优配置策略可以通过求解以下最优问题来得到:

$$\tau^{N*}_{\mathrm{IDS}_i}(u) := \beta^*_{\mathrm{IDS}_i} = \underset{\beta_{\mathrm{IDS}_i} \in \Delta(A_{\mathrm{IDS}})}{\operatorname{argmax}} J^N_{\mathrm{IDS}_i}(u_0, \tau^N_{\mathrm{IDS}_i}(u)) \tag{15}$$

$$\text{s.t.} \quad \pi^N_{\gamma_i}(b) = \pi^{N*}_{\gamma_i}(b) \tag{16}$$

其中, $\pi^{N*}_{\gamma_i}(b)$ 是满足式(11)和(12)的攻击者的稳态纳什均衡策略。

2.3 检测库分配

在检测库分配过程中,会有一个入侵检测系统充当中心管理者调度所有入侵检测系统加载检测库。在得到各入侵检测系统的最优配置策略 $\{\tau^{N*}_{\mathrm{IDS}_1}, \tau^{N*}_{\mathrm{IDS}_2}, \cdots, \tau^{N*}_{\mathrm{IDS}_N}\}$ 以及对应攻击者的稳态纳什均衡策略 $\{\pi^{N*}_{\gamma_1}, \pi^{N*}_{\gamma_2}, \cdots, \pi^{N*}_{\gamma_N}\}$ 后,将该信息报告给入侵检测网络管理者,管理者通过基于策略共享的检测库分配方法得到协作式入侵检测系统实际的检测库配置方案为 $\{a_{\mathrm{IDS}_1}, a_{\mathrm{IDS}_2}, \cdots, a_{\mathrm{IDS}_N}\}$, 该分配方法将在下一节中详细介绍。此时入侵检测系统 IDS_i 的期望收益为

$$\tilde{r}_{\mathrm{IDS}_i}(u, \{a_{\mathrm{IDS}_i}, \pi^{N*}_{\gamma_i}\}) = \sum_{a_{\gamma_i} \in A_\gamma} \pi^{N*}_{\gamma_i}(a_{\gamma_i}) \omega(s, a_{\mathrm{IDS}_i}, a_{\gamma_i}) \tag{17}$$

其中, $u = \{s, b\}$ 表示当前时刻子系统 i 的状态以及攻击者 γ_i 的信念, $\pi^{N*}_{\gamma_i}$ 是攻击者 γ_i 的稳态纳什均衡

策略，$\omega(s, a_{\mathrm{IDS}_i}, a_{\gamma_i})$ 已在式(2)定义。

此外，本文考虑的是入侵检测网络的收益，即所有入侵检测系统的收益之和，入侵检测网络的收益为

$$r_{\mathrm{IDN}} = \sum_{i=1}^{N} \tilde{r}_{\mathrm{IDS}_i}(u, \{a_{\mathrm{IDS}_i}, \pi_{\gamma_i}^{N*}\}) \tag{18}$$

因此，中心管理者求得的检测库配置方案需满足：

$$\max_{a_{\mathrm{IDS}_i}} r_{\mathrm{IDN}}$$

$$\text{s.t.} \quad a_{\mathrm{IDS}_1} \cap a_{\mathrm{IDS}_2} \cap \cdots \cap a_{\mathrm{IDS}_N} = \varnothing$$

$$a_{\mathrm{IDS}_1} \cup a_{\mathrm{IDS}_2} \cup \cdots \cup a_{\mathrm{IDS}_N} \subseteq L$$

3 入侵检测系统检测库配置方案计算

本节将给出具体求解基于信念的随机博弈、混合 Markov 决策过程以及检测库分配的方法。

3.1 攻击策略

为了解决有限时间非完全信息博弈问题 1，本文提出后向递归算法来构建攻击者 γ_i 的最优稳态策略。在该算法中，假定加权入侵检测系统 $\overline{IDS_i}$ 已经使用稳态纳什均衡策略；并且证明该算法的解就是有限时间非完全信息博弈问题 G_N 的稳态纳什均衡策略。

令 G_N 中的 SNE 策略为 $(\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b))$，具体可以表示为 $\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b) = (\pi_{\overline{\mathrm{IDS}_i},0}^{N*}(b_0), \cdots, \pi_{\overline{\mathrm{IDS}_i},N}^{N*}(b_N))$ 和 $\pi_{\gamma_i}^{N*}(b) = (\pi_{\gamma_i,0}^{N*}(b_0), \cdots, \pi_{\gamma_i,N}^{N*}(b_N))$。根据式(10)，在稳态纳什均衡策略下，攻击者 γ_i 从 t 时刻到 N 时刻的累计收益函数为

$$J_{\gamma_i,t}^{N}(b_t, \{\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b)\}) = \sum_{t}^{N} r_{\gamma_i}(b_t, \{\pi_{\overline{\mathrm{IDS}_i},t}^{N*}(b_t), \pi_{\gamma_i,t}^{N*}(b_t)\}) \tag{19}$$

其中，b_t 表示 t 时刻攻击者 γ_i 的信念，$\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b)$ 和 $\pi_{\gamma_i}^{N*}(b)$ 分别表示加权入侵检测系统 $\overline{IDS_i}$ 和攻击者 γ_i 与信念 b 有关的稳态纳什均衡策略。

下面给出后向递归算法的具体步骤。

(1)初始化：对所有的信念 $b \in B$，令 $t = N+1$ 时攻击者 γ_i 的收益函数为

$$J_{\gamma_i,N+1}^{N}(b) = 0 \tag{20}$$

则 $t = N$ 时刻攻击者 γ_i 的最优稳态策略为

$$\pi_{\gamma_i,N}^{N*}(b_N) := \beta_{\gamma_i}^{*} = \underset{a_{\gamma_i} \in \Delta(A_\gamma)}{\operatorname{argmax}} r_{\gamma_i}(b_N, \pi_{\overline{\mathrm{IDS}_i},N}^{N*}(b_N), \beta_{\gamma_i}) \tag{21}$$

相应地，$t = N$ 时攻击者 γ_i 的收益函数为

$$J_{\gamma_i,N}^{N}(b_N) = r_{\gamma_i}(b_N, \pi_{\overline{\mathrm{IDS}_i},N}^{N*}(b_N), \pi_{\gamma_i,N}^{N*}(b_N)) \tag{22}$$

(2)递归：对于任意 $t = N-1, N-2, \cdots, 0$；$\forall b_t \in B$，攻击者 γ_i 的最优稳态策略为

$$\pi_{\gamma_i,t}^{N*}(b_t) := \beta_{\gamma_i}^{*} = \underset{a_{\gamma_i} \in \Delta(A_\gamma)}{\operatorname{argmax}} \Big[r_{\gamma_i}(b_t, \pi_{\overline{\mathrm{IDS}_i},t}^{N*}(b_t), \beta_{\gamma_i}) + \sum_{b'} T(b' \mid b_t, \pi_{\overline{\mathrm{IDS}_i},t}^{N*}(b_t), \beta_{\gamma_i}) J_{\gamma_i,t+1}^{N}(b') \Big] \tag{23}$$

则相应地，攻击者 γ_i 的收益函数为

$$J_{\gamma_i,t}^{N}(b_t) = r_{\gamma_i}(b_t, \pi_{\overline{\mathrm{IDS}_i},t}^{N*}(b_t), \pi_{\gamma_i,t}^{N*}(b_t)) + \sum_{b'} T(b' \mid b_t, \pi_{\overline{\mathrm{IDS}_i},t}^{N*}(b_t), \pi_{\gamma_i,t}^{N*}(b_t)) J_{\gamma_i,t+1}^{N}(b') \tag{24}$$

(3)构造攻击者 γ_i 的一个最优稳态策略为

$$\pi_{\gamma_i}^{N*}(b) = (\pi_{\gamma_i,0}^{N*}(b_0), \pi_{\gamma_i,1}^{N*}(b_1), \cdots, \pi_{\gamma_i,N}^{N*}(b_N))$$

在上述的后向递归算法中，理论上需要先将一方参与者的稳态纳什均衡策略固定下来，再求解另一方参与者的稳态纳什均衡策略，但是在实际的仿真案例中，由于 G_N 是零和博弈，因此在求得收益矩阵以后可以通过效用等值[22]方法同时求解出双方的稳态纳什均衡策略。下面的定理 1 给出了上述后向递归算法与有限时间非完全信息博弈 G_N 中稳态纳什均衡存在的关系。

定理 1 由后向递归算法构造出的解 $(\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b))$ 是有限时间非完全信息博弈 G_N 的一个稳态纳什均衡策略。

证明 要证明 $(\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b))$ 是有限时间非完全信息博弈 G_N 的一个稳态纳什均衡策略，根据问题 1 中对稳态纳什均衡的定义，即等价于证明 $J_{\overline{\mathrm{IDS}_i}}^{N}(b_0, \{\pi_{\overline{\mathrm{IDS}_i}}^{N}(b), \pi_{\gamma_i}^{N*}(b)\}) \leqslant J_{\overline{\mathrm{IDS}_i}}^{N}(b_0, \{\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b)\})$ $J_{\gamma_i}^{N}(b_0, \{\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N}(b)\}) \leqslant J_{\gamma_i}^{N}(b_0, \{\pi_{\overline{\mathrm{IDS}_i}}^{N*}(b), \pi_{\gamma_i}^{N*}(b)\})$ 成立。由于在后向递归算法中假定加权入侵检测系统 $\overline{IDS_i}$ 使用稳态纳什均衡策

略 $\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b)$,所以,$J^N_{\overline{\mathrm{IDS}_i}}(b_0,\{\pi^N_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\}) \leqslant J^N_{\overline{\mathrm{IDS}_i}}(b_0,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})$ 显然是成立的。因此,下面只需要证明 $J^N_{\gamma_i}(b_0,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^N_{\gamma_i}(b)\}) \leqslant J^N_{\gamma_i}(b_0,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})$ 成立。

首先,要证明当 $t=N$ 时,不等式 $J^N_{\gamma_i,N}(b_N,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^N_{\gamma_i}(b)\}) \leqslant J^N_{\gamma_i,N}(b_N,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})$ 成立。根据式(21)和(22),可得到:

$$
\begin{aligned}
&J^N_{\gamma_i,N}(b_N,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})\\
&=\max_{\beta_{\gamma_i}\in\Delta(A_\gamma)} r_{\gamma_i}(b_N,\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\beta_{\gamma_i})\\
&\geqslant r_{\gamma_i}(b_N,\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\beta_{\gamma_i})=J^N_{\gamma_i,N}(b,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^N_{\gamma_i}(b)\})
\end{aligned}
$$

接着,当 $t=N-1,N-2,\cdots,0$ 时,同样可以得到 $J^N_{\gamma_i,t}(b_t,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})$

$$
\begin{aligned}
&=r_{\gamma_i}(b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\pi^{N*}_{\gamma_i,t}(b))\\
&\quad+\sum_{b'}T(b'\mid b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\pi^{N*}_{\gamma_i,t}(b))J^N_{\gamma_i,t+1}(b')\\
&=\max_{\beta_{\gamma_i}\in\Delta(A_\gamma)}\Big[r_{\gamma_i}(b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\beta_{\gamma_i})\\
&\quad+\sum_{b'}T(b'\mid b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\beta_{\gamma_i})J^N_{\gamma_i,t+1}(b')\Big]\\
&\geqslant r_{\gamma_i}(b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\beta_{\gamma_i})\\
&\quad+\sum_{b'}T(b'\mid b_t,\pi^{N*}_{\overline{\mathrm{IDS}_i},t}(b),\beta_{\gamma_i})J^N_{\gamma_i,t+1}(b')\\
&=J^N_{\gamma_i,t}(b_t,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^N_{\gamma_i}(b)\})
\end{aligned}
$$

其中,第2个等号成立的条件是用式(23)求解出攻击者的最优稳态策略。因此,要证的不等式 $J^N_{\gamma_i}(b_0,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^N_{\gamma_i}(b)\}) \leqslant J^N_{\gamma_i}(b_0,\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\})$ 总是成立的,即证明了有限时间非完全信息博弈 G_N 中稳态纳什均衡的存在。

值得注意的是,在定义有限时间非完全信息博弈的收益函数时没有用到折扣因子(或者,等价于折扣因子 $\rho=1$),如果在定义有限时间非完全信息博弈的收益函数时加上折扣因子,定理1也依然适用于证明该博弈存在一个稳态纳什均衡,详见下面的推论1。

推论1 当有限时间非完全信息博弈 G_N 的收益函数定义为

$$J^N_{\overline{\mathrm{IDS}_i}}(b_0,\pi^N(b))=\sum_{t=0}^{N}\rho^t r_{\overline{\mathrm{IDS}_i}}(b,\pi^N(b)) \tag{25}$$

$$J^N_{\gamma_i}(b_0,\pi^N(b))=\sum_{t=0}^{N}\rho^t r_{\gamma_i}(b,\pi^N(b)) \tag{26}$$

其中,$0<\rho<1$ 是折扣因子。在这样的情况下,G_N 依然存在一个稳态纳什均衡。

证明 在定义式(19)中加入折扣因子 ρ,后续证明过程和定理1的证明过程相同。

3.2 检测库配置方案

在这一小节给出求解混合 Markov 决策过程的方法。定理2给出了入侵检测系统 IDS_i 的最优纯动作贝尔曼方程。

定理2 入侵检测系统 IDS_i 的最优纯动作贝尔曼方程为

$$
\begin{aligned}
J^{N*}_{\mathrm{IDS}_i}(u)=&\max_{\beta_{\mathrm{IDS}_i}\in\Delta(A_{\mathrm{IDS}})}\sum_a\beta_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i})\\
&\times Q^{N*}_{\mathrm{IDS}_i}(u,a)
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
Q^{N*}_{\mathrm{IDS}_i}(u,a)=&\bar r_{\mathrm{IDS}_i}(u,a)+\rho\sum_{s'\in S}\sum_{b'\in B}T_{s,s'}(a_{\mathrm{IDS}_i},a_{\gamma_i})\\
&\times T(b'\mid b,\{\beta^*_{\overline{\mathrm{IDS}_i}},\beta^*_{\gamma_i}\},\bar a)J^{N*}_{\mathrm{IDS}_i}(u')
\end{aligned} \tag{28}
$$

其中,$\beta^*_{\overline{\mathrm{IDS}_i}}=\pi^*_{\overline{\mathrm{IDS}_i}}(b)$,$\bar r_{\mathrm{IDS}_i}(u,a)=\omega(s,a_{\mathrm{IDS}_i},a_{\gamma_i})$,$T(b'\mid b,\{\beta^*_{\overline{\mathrm{IDS}_i}},\beta^*_{\gamma_i}\},\bar a)=\Pr(a^o_t=\bar a\mid B_t=b,a_t=\bar a)$,当 $b'=b\boldsymbol{T}(a_{\overline{\mathrm{IDS}_i}},a_{\gamma_i})$ 时,$\bar a=\{a_{\overline{\mathrm{IDS}_i}},a_{\gamma_i}\}$,$a_{\overline{\mathrm{IDS}_i}}$ 是入侵检测系统 IDS_i 根据策略 $\beta^*_{\overline{\mathrm{IDS}_i}}$ 采取的欺骗动作。

证明 根据基于信念的随机博弈中求得的稳态纳什均衡策略 $\pi^{N*}(b)=\{\pi^{N*}_{\overline{\mathrm{IDS}_i}}(b),\pi^{N*}_{\gamma_i}(b)\}$,$\forall b\in B$,入侵检测系统 IDS_i 的累计收益函数满足最优贝尔曼方程[23]:

$$
\begin{aligned}
J^{N*}_{\mathrm{IDS}_i}(u)=&\max_{\beta_{\mathrm{IDS}_i}\in\Delta(A_{\mathrm{IDS}})}Q^{N*}_{\mathrm{IDS}_i}(u,\{\beta_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
Q^{N*}_{\mathrm{IDS}_i}(u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})=&r_{\mathrm{IDS}_i}(u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})\\
&+\rho\sum_{u'\in U}\hat T(u'\mid u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})J^{N*}_{\mathrm{IDS}_i}(u')
\end{aligned} \tag{30}
$$

其中,$\beta^*_{\gamma_i}=\pi^{N*}_{\gamma_i}(b)$。再由混合状态转移概率的定义可得:

$$
\begin{aligned}
&\hat T(u'\mid u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})\\
&=\Pr(U_{t+1}=\{s',b'\}\mid U_t=\{s,b\},\hat A_t=\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\})\\
&=\Pr\{B_{t+1}=b'\mid B_t=b,\hat A_t=\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\}\}\\
&\quad\times\Pr\{S_{t+1}=s'\mid S_t=s,\hat A_t=\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\}\}\\
&=\beta^*_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i})\boldsymbol{T}_{s,s'}(a_{\mathrm{IDS}_i},a_{\gamma_i})T(b'\mid b,\{\beta^*_{\overline{\mathrm{IDS}_i}},\\
&\quad\beta^*_{\gamma_i}\})
\end{aligned}
$$

其中,第 2 个等号成立的条件是因为状态 s 的转移与信念 b 的转移都是相互独立的;第 3 个等号成立的条件是因为攻击者 γ_i 无法获知入侵检测系统 IDS_i 的策略 β_{IDS_i},所以只能根据加权入侵检测系统 $\overline{IDS_i}$ 的稳态纳什均衡策略 $\beta^*_{\overline{\mathrm{IDS}_i}}$ 来更新信念。因此,式(30)可以改写为

$$Q^{N*}_{\mathrm{IDS}_i}(u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\}) = r_{\mathrm{IDS}_i}(u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\}) + \rho\sum_{s'\in S}\sum_{b'\in B}\beta^*_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i}) \times \boldsymbol{T}_{s,s'}(a_{\mathrm{IDS}_i},a_{\gamma_i})T(b'\mid b,\{\beta^*_{\overline{\mathrm{IDS}_i}},\beta^*_{\gamma_i}\})J^{N*}_{\mathrm{IDS}_i}(u') \tag{31}$$

接下来,令 $Q^{N*}_{\mathrm{IDS}_i}(u,a)$ 表示入侵检测系统 IDS_i 的最优纯动作状态值函数,其中,$a=\{a_{\mathrm{IDS}_i},a_{\gamma_i}\}$,$a_{\mathrm{IDS}_i}$ 是入侵检测系统 IDS_i 根据最优配置策略 $\beta^*_{\mathrm{IDS}_i}$ 选择的纯动作,a_{γ_i} 是攻击者 γ_i 根据稳态纳什均衡策略 $\beta^*_{\gamma_i}$ 选择的纯动作。由于双方动作是相互独立的,有 $\beta^*(a)=\beta^*_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i})$,因此可以得到:

$$Q^{N*}_{\mathrm{IDS}_i}(u,\{\beta^*_{\mathrm{IDS}_i},\beta^*_{\gamma_i}\}) = \sum_a \beta^*_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i})Q^{N*}_{\mathrm{IDS}_i}(u,a)$$

则式(29)可以改写为

$$J^{N*}_{\mathrm{IDS}_i}(u) = \max_{\beta_{\mathrm{IDS}_i}\in\Delta(A_{\mathrm{IDS}})}\sum_a \beta^*_{\mathrm{IDS}_i}(a_{\mathrm{IDS}_i})\beta^*_{\gamma_i}(a_{\gamma_i})Q^{N*}_{\mathrm{IDS}_i}(u,a)$$

将式(31)中的概率动作换成纯动作,可得入侵检测系统 IDS_i 的最优纯动作状态值函数为

$$Q^{N*}_{\mathrm{IDS}_i}(u,a) = \bar{r}_{\mathrm{IDS}_i}(u,a) + \rho\sum_{s'\in S}\sum_{b'\in B}T_{s,s'}(a_{\mathrm{IDS}_i},a_{\gamma_i}) \times T(b'\mid b,\{\beta^*_{\overline{\mathrm{IDS}_i}},\beta^*_{\gamma_i}\},\bar{a})J^{N*}_{\mathrm{IDS}_i}(u')$$

其中,$a=\{a_{\mathrm{IDS}_i},a_{\gamma_i}\}$。

综上所述,入侵检测系统 IDS_i 的最优纯动作贝尔曼方程式(27)和(28)得证。

通过最优纯动作贝尔曼方程式(27)和(28),入侵检测系统 IDS_i 可以算得任意状态下所有联合动作 $a=\{a_{\mathrm{IDS}_i},a_{\gamma_i}\}$ 对应的 $Q^{N*}_{\mathrm{IDS}_i}(u,a)$,即状态纯动作值表。又因为入侵检测系统在信息方面有优势,可以看到攻击者的策略制定过程,即在博弈的每一阶段,入侵检测系统都能获得攻击者的稳态纳什均衡策略 $\beta^*_{\gamma_i}$。因此,入侵检测系统 IDS_i 可以通过算法 1求得最优检测库配置方案。

算法 1 入侵检测系统最优检测库配置方案

1. 初始化系统状态 s 和信念状态 b
2. 攻击者 γ_i 通过式(23)算得稳态纳什均衡策略 $\beta^*_{\gamma_i}$
3. 入侵检测系统 IDS_i 观察到攻击者 γ_i 稳态纳什均衡策略 $\beta^*_{\gamma_i}$,并通过式(28)算得当前状态下所有联合纯动作 $\{a_{\mathrm{IDS}_i},a_{\gamma_i}\}$ 对应的值 $Q^{N*}_{\mathrm{IDS}_i}(u,\{a_{\mathrm{IDS}_i},a_{\gamma_i}\})$
4. **for** $j=\{1,2,\cdots,F_M\}$ **do**
5. $\eta(j)=\sum_{z\in A_\gamma}\beta^*_{\gamma_i}(z)Q^{N*}_{\mathrm{IDS}_i}(u,\{j,z\})$
6. **End for**
7. $a_{\mathrm{IDS}_i}=\arg\max_j\{\eta(j)\}$
8. 输出入侵检测系统 IDS_i 的配置方案 a_{IDS_i}

算法 1 的步骤 4 表示遍历入侵检测系统 IDS_i 的检测库配置方案 A_{IDS};步骤 5 表示入侵检测系统 IDS_i 每一个检测库配置方案对应的期望收益,$\beta^*_{\gamma_i}(z)$ 表示攻击者选择 z 类型的攻击的概率;步骤7 表示选出能使期望收益最大的检测库配置方案。

由该算法流程可以看出,入侵检测系统 IDS_i 的最优检测库配置方案是在获得防守者 γ_i 的稳态纳什均衡策略 $\beta^*_{\gamma_i}$ 以后,计算每一种检测库配置方案对应的期望收益,通过比较期望收益的大小,选择最大值对应的配置方案。因此,入侵检测系统 IDS_i 的最优检测库配置方案实质上是一个纯策略。

3.3 检测库分配

由于协作式入侵检测系统共用一套检测库,因此可能会出现同一时刻多个入侵检测系统加载同一检测库的矛盾,本小节将介绍如何用基于策略共享的集中式分配方法解决该矛盾。

每一时刻,入侵检测系统通过与攻击者的交互得到最优检测库配置方案。由算法 1 的分析可知,该最优检测库配置方案是纯策略,表示为 $\{a_{\mathrm{IDS}_1},\cdots,a_{\mathrm{IDS}_i},\cdots,a_{\mathrm{IDS}_N}\}$,其中,$a_{\mathrm{IDS}_i}\in A_{\mathrm{IDS}}$,$i\in\{1,2,\cdots,N\}$。将该最优检测库配置方案报告给中心管理者,若满足式(4)和(5),则说明各入侵检测系统的检测库配置方案没有任何矛盾,即没有 2 个或 2 个以上的入侵检测系统同时加载同一个检测库,则中心管理者按该方案分配检测库。若不满足上述条件则说明有检测库被重复加载,则此时中心管理者先将没有矛盾的检测库分配掉;然后再通过一个优化机制将矛盾的检测库分配给合适的入侵检测系统。该优

化机制具体描述如下。

(1)对任意入侵检测系统 IDS_i ($i \in \{1,2,\cdots,N\}$),找出其与剩余所有入侵检测系统加载有矛盾的检测库的集合,表示为 $u_{ij} = a_{\mathrm{IDS}_i} \cap a_{\mathrm{IDS}_j}$ ($j = \{1,2,\cdots i-1, i+1,\cdots,N\}$)。

(2)对 $\forall l_m \in u_{ij}$,利用式(18)计算入侵检测系统 IDS_i 加载 l_m、IDS_j 不加载 l_m 后入侵检测网络的期望收益 r^i_{IDN}。注意,当从入侵检测系统 IDS_j 的配置方案中去掉检测库 l_m 以后,中心管理者会再从剩余的检测库中挑选满足以下条件的检测库分配给 IDS_j。

$$\max_{l_h \in L \setminus \{l_m\}} \Big[\sum_{a_{\gamma_j} \in A_\gamma} \beta^*_{\gamma_j}(a_{\gamma_j}) p_{l_h, a_{\gamma_j}} D(s_j, a_{\gamma_j}) + \delta_{\mathrm{IDS}_j} \bar{c}_{\mathrm{IDS}_j}(l_h) - \delta_{\gamma_j} c_{\gamma_j}(a_{\gamma_j}) \Big]$$

再用相同的方法计算入侵检测系统 IDS_j 加载 l_m、IDS_i 不加载 l_m 后入侵检测网络的期望收益 r^j_{IDN}。

(3)比较 r^i_{IDN} 和 r^j_{IDN} 的大小,当 $r^i_{\mathrm{IDN}} > r^j_{\mathrm{IDN}}$ 时,入侵检测系统 IDS_i 获得检测库 l_m;当 $r^i_{\mathrm{IDN}} < r^j_{\mathrm{IDN}}$ 时,入侵检测系统 IDS_j 获得检测库 l_m;当 $r^i_{\mathrm{IDN}} = r^j_{\mathrm{IDN}}$ 时,检测库 l_m 可以被任意分配给入侵检测系统 IDS_i 或 IDS_j。

4 仿真分析

如图 2 所示,考虑一个由 2 个入侵检测系统相互连接的入侵检测网络展开有限次博弈,每个子系统 i($i \in \{1,2\}$) 都有 3 个状态 $\{S_1, S_2, S_3\}$。攻击者的纯动作集合为 $A_\gamma = \{1,2\}$,且不同类型的攻击的能耗分别为 $c_{\gamma_i}(1) = 0.8$ 和 $c_{\gamma_i}(2) = 1.2$ ($i \in \{1,2\}$)。不同状态下各类型的攻击对子系统 i 造成的损失如表 1 所示。

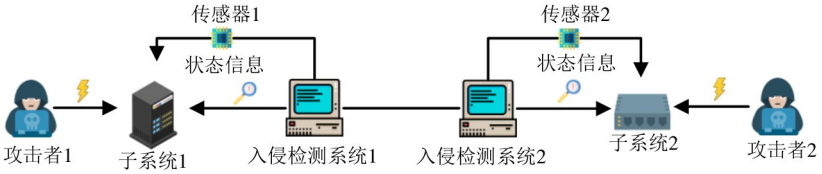

图 2 二对二攻防博弈模型图

表 1 不同状态下攻击对系统造成的损失

状态 \ 攻击类型	$a_{\gamma_i} = 1.0$	$a_{\gamma_i} = 2.0$
S_1	1.5	2.0
S_2	2.5	2.0
S_3	3.0	1.0

协作式入侵检测系统共用的检测库集合为 $L = \{l_1, l_2\}$,每个入侵检测系统任意时刻只能加载一个检测库,则入侵检测系统的纯动作(检测库配置方案)集合为 $A_{\mathrm{IDS}} = \{\{l_1\},\{l_2\}\}$,每个检测库被加载的能耗分别为 $\bar{c}_{\mathrm{IDS}_i}(l_1) = 0.6, \bar{c}_{\mathrm{IDS}_i}(l_2) = 1.0, i \in \{1,2\}$。各检测库成功检测到攻击的概率分别为 $p_{l_1,1} = 0.80, p_{l_1,2} = 0.20, p_{l_2,1} = 0.15, p_{l_2,2} = 0.90$。式(2)中的权重参数 $\delta_{\mathrm{IDS}_i} = 0.5$ 和 $\delta_{\gamma_i} = 0.4, i \in \{1,2\}$。状态转移矩阵 $\boldsymbol{T}(a_{\mathrm{IDS}_i}, a_{\gamma_i})$ 可以按以下方式计算得到。假设当前子系统 i 的状态为 S_2,攻击者 γ_i 发动类型为 1 的攻击,对应的 IDS_i 加载检测库 $\{l_1\}$,则相关的状态转移概率分别为

$$\boldsymbol{T}_{2,3}(\{l_1\},1) = 1 - p_{l_1,1} = 0.2$$

$$\boldsymbol{T}_{2,1}(\{l_1\},1) = p_{l_1,1} = 0.8, \boldsymbol{T}_{2,2}(\{l_1\},1) = 0.0$$

用相同的方法可以得到不同状态所有联合动作下的状态转移概率。令 IDS_i 在状态 S_k($k = 1,2,3$) 下的收益矩阵为 $\boldsymbol{\omega}(k, \cdot, \cdot)$,则 $\boldsymbol{\omega}(1, \cdot, \cdot) = \begin{bmatrix} 0.28 & 1.42 \\ 1.455 & 0.22 \end{bmatrix}$, $\boldsymbol{\omega}(2, \cdot, \cdot) = \begin{bmatrix} 0.48 & 1.42 \\ 2.305 & 0.22 \end{bmatrix}$ 和 $\boldsymbol{\omega}(3, \cdot, \cdot) = \begin{bmatrix} 0.58 & 0.62 \\ 2.73 & 0.12 \end{bmatrix}$。

通过后向递归算法和算法 1 可以求解有限时间非完全信息博弈的稳态纳什均衡策略和最优检测库配置策略。将博弈的长度设置为 $N = 3$。当 $t = 3$ 时,信念可以表示为 $b_3 = (b_3(1), b_3(2), b_3(3))$, $b_3(1) + b_3(2) + b_3(3) = 1$,且 $b_3(1) > 0, b_3(2) > 0, b_3(3) > 0$。令 $\boldsymbol{\pi}^{N*}_{\gamma_i,3}(b_3)$ 和 $\boldsymbol{\pi}^{N*}_{\mathrm{IDS}_i,3}(b_3)$ 分别表示

博弈第 3 阶段攻击者 γ_i 和加权入侵检测系统 $\overline{IDS_i}$ 的稳态纳什均衡策略，$a_{\mathrm{IDS}_i,3}$ 表示博弈第 3 阶段入侵检测系统 IDS_i 的最优检测库配置方案。根据后向递归算法，首先对 $\forall b \in B$，设置 $J^N_{\gamma_i,4}(b)=0$；然后根据式(21)可以算得 $\pi^{N*}_{\gamma_i,3}(b_3)$ 和 $\pi^{N*}_{\overline{\mathrm{IDS}_i},3}(b_3)$ 为

$$(\pi^{N*}_{\gamma_i,3}(1\mid b_3),\pi^{N*}_{\gamma_i,3}(2\mid b_3))=(q_4-q_2/\epsilon,\ q_1-q_3/\epsilon)$$

$$(\pi^{N*}_{\overline{\mathrm{IDS}_i},3}(\{l_1\}\mid b_3),\pi^{N*}_{\overline{\mathrm{IDS}_i},3}(\{l_2\}\mid b_3))=(q_4-q_3/\epsilon,\ q_1-q_2/\epsilon)$$

其中，$\epsilon=q_1-q_2-q_3+q_4$，且 $q_m(m\in\{1,\cdots,4\})$ 是第 3 阶段博弈矩阵的元素，其具体表达式为

$$q_1=0.280b_3(1)+0.480b_3(2)+0.580b_3(3)$$

$$q_2=1.420b_3(1)+1.420b_3(2)+0.620b_3(3)$$

$$q_3=1.455b_3(1)+2.305b_3(2)+2.730b_3(3)$$

$$q_4=0.220b_3(1)+0.220b_3(2)+0.120b_3(3)$$

最后，利用式(22)可以求得 $J^N_{\gamma_i,3}(b_3)$。根据算法 1 可求得入侵检测系统 IDS_i 分别加载检测库 l_1 和 l_2 的收益的差为

$$\begin{aligned}\eta&=J_{\mathrm{IDS}_i}(u_3\mid a_{\mathrm{IDS}_i,3}=\{l_1\})-J_{\mathrm{IDS}_i}(u_3\mid a_{\mathrm{IDS}_i,3}=\{l_2\})\\&=\pi^{N*}_{\gamma_i,3}(1\mid b_3)[\omega(m,\{l_1\},1)-\omega(m,\{l_2\},1)]\\&\quad+\pi^{N*}_{\gamma_i,3}(2\mid b_3)[\omega(m,\{l_1\},2)-\omega(m,\{l_2\},2)]\end{aligned}$$

当 $\eta>0$ 时，入侵检测系统 IDS_i 加载检测库 l_1；当 $\eta<0$ 时，入侵检测系统 IDS_i 加载检测库 l_2；当 $\eta=0$ 时，任意加载检测库 l_1 或者检测库 l_2。在得到最优检测库配置方案 $\{a_{\mathrm{IDS}_1},a_{\mathrm{IDS}_2}\}$ 后，如果满足式(4)和(5)，则直接为 IDS_1 和 IDS_2 分配检测库；否则按 3.3 节中的分配方法为 IDS_1 和 IDS_2 分配检测库。

在博弈的 $t=2$、1、0 阶段，信念 b_t 可以用信念更新式(6)得到，真实状态通过状态转移矩阵得到，攻击者的稳态纳什均衡策略和入侵检测系统的最优配置方案可通过上述相同的步骤算得。

在这个仿真案例中，给定攻击者 γ_1 的信念为 $b_3=(0.5,0.3,0.2)$，子系统 1 的真实状态为 1；攻击者 γ_2 的信念为 $b_3=(0.4,0.4,0.2)$，子系统 2 的真实状态为 2，按上述步骤，攻击者 γ_1 从博弈的第 3 阶段到第 0 阶段的稳态纳什均衡策略分别为

$$(\pi^{N*}_{\gamma_1,3}(1\mid b_3),\pi^{N*}_{\gamma_1,3}(2\mid b_3))=(0.40,0.60)$$

$$(\pi^{N*}_{\gamma_1,2}(1\mid b_2),\pi^{N*}_{\gamma_1,2}(2\mid b_2))=(0.44,0.56)$$

$$(\pi^{N*}_{\gamma_1,1}(1\mid b_1),\pi^{N*}_{\gamma_1,1}(2\mid b_1))=(0.35,0.65)$$

$$(\pi^{N*}_{\gamma_1,0}(1\mid b_0),\pi^{N*}_{\gamma_1,0}(2\mid b_0))=(0.27,0.73)$$

相对应地，入侵检测系统 IDS_1 从博弈的第 3 阶段到第 0 阶段的最优检测库配置方案为 $a_{\mathrm{IDS}_1,3}=\{l_1\}$，$a_{\mathrm{IDS}_1,2}=\{l_2\}$，$a_{\mathrm{IDS}_1,1}=\{l_2\}$ 和 $a_{\mathrm{IDS}_1,0}=\{l_2\}$。

用同样的方法可以得到攻击者 γ_2 从第 3 阶段到第 0 阶段的稳态纳什均衡策略分别为

$$(\pi^{N*}_{\gamma_2,3}(1\mid b_3),\pi^{N*}_{\gamma_2,3}(2\mid b_3))=(0.39,0.61)$$

$$(\pi^{N*}_{\gamma_2,2}(1\mid b_2),\pi^{N*}_{\gamma_2,2}(2\mid b_2))=(0.31,0.69)$$

$$(\pi^{N*}_{\gamma_2,1}(1\mid b_1),\pi^{N*}_{\gamma_2,1}(2\mid b_1))=(0.40,0.60)$$

$$(\pi^{N*}_{\gamma_2,0}(1\mid b_0),\pi^{N*}_{\gamma_2,0}(2\mid b_0))=(0.46,0.54)$$

相对应地，入侵检测系统 IDS_2 从博弈的第 3 阶段到第 0 阶段的最优检测库配置方案 $a_{\mathrm{IDS}_2,3}=\{l_1\}$，$a_{\mathrm{IDS}_2,2}=\{l_1\}$，$a_{\mathrm{IDS}_2,1}=\{l_2\}$ 和 $a_{\mathrm{IDS}_2,0}=\{l_2\}$。

注意到在博弈的第 3 阶段，入侵检测系统 IDS_1 和 IDS_2 都要加载检测库 l_1。当入侵检测系统 IDS_1 加载检测库 l_1 而 IDS_2 加载检测库 l_2，入侵检测网络的收益 $r^1_{\mathrm{IDN}}=2.585$；当入侵检测系统 IDS_2 加载检测库 l_1 而 IDS_1 加载检测库 l_2，入侵检测网络的收益 $r^2_{\mathrm{IDN}}=1.935$。由于 $r^1_{\mathrm{IDN}}>r^2_{\mathrm{IDN}}$，因此把检测库 l_1 分配给入侵检测系统 IDS_1。类似地在博弈的第 1 阶段，入侵检测系统 IDS_1 和 IDS_2 都要加载检测库 l_2。当入侵检测系统 IDS_1 加载检测库 l_2 而 IDS_2 加载检测库 l_1，入侵检测网络的收益 $r^1_{\mathrm{IDN}}=0.740$；当入侵检测系统 IDS_2 加载检测库 l_2 而 IDS_1 加载检测库 l_1，入侵检测网络的收益 $r^2_{\mathrm{IDN}}=0.740$，由于 $r^1_{\mathrm{IDN}}=r^2_{\mathrm{IDN}}$，因此检测库 l_2 可以分配给入侵检测系统 IDS_1 或者 IDS_2。在博弈的第 0 阶段，入侵检测系统 IDS_1 和 IDS_2 都要加载检测库 l_2。当入侵检测系统 IDS_1 加载检测库 l_2 而 IDS_2 加载检测库 l_1，入侵检测网络的收益 $r^1_{\mathrm{IDN}}=4.150$；当入侵检测系统 IDS_2 加载检测库 l_2 而 IDS_1 加载检测库 l_1，入侵检测网络的收益 $r^2_{\mathrm{IDN}}=0.800$，由于 $r^1_{\mathrm{IDN}}>r^2_{\mathrm{IDN}}$，因此检测库 l_2 被分配给入侵检测系统 IDS_1。

在相同模拟场景下比较了本文提出的基于策略共享的集中式分配方法和文献[14]提出的分布式激励分配方法解决检测库重复加载的矛盾后，所有入侵检测系统的收益总和。前者为 10.360，后者为 9.235，由此可以看出，本文提出的分配方法比分布

式激励分配方法在收益值上提升了12.18%。此外,还研究了2种分配方法完成分配的时间,基于策略共享的集中式分配方法运行时长为2.3 ms,分布式激励分配方法运行时长为1.4 ms,后者的运行时长相对减少了39.00%。因此当分配次数变多以及检测库数量变多后,分布式激励分配方法在运行时间上的优势会更明显。这是由于本文提出的分配方法为了能最优化整体入侵检测网络的检测性能,考虑了矛盾检测库分配的所有情况,因此当检测库的数量较多时耗时会较为严重。而分布式激励分配方法是各检测系统选择能使自己检测性能最优化的检测库,即通过一系列的局部最优来近似整体最优,这种分配方法能降低计算的复杂度,但是整个入侵检测系统的性能可能不会达到最优。

5 结 论

本文给出了有限时间下非完全信息协作式入侵检测系统的最优检测库配置方法,该方法不仅解决了各入侵检测系统应对不同类型攻击时最优检测库的配置,而且解决了共用一套检测库引发的加载库的矛盾。攻击者无法获知系统状态的设定也更符合实际,可以通过构建一个基于信念的随机博弈来解决,并借助后向递归算法证明了博弈中SNE的存在。真正的入侵检测系统利用一个欺骗机制获知攻击者的策略,并基于该策略算得每种检测库配置方案下的期望收益,通过比较期望收益的大小得到最优检测库配置方案。解决多个入侵检测系统加载同一检测库矛盾的分配方法实质上就是在分配完所有检测库的约束下最大化入侵检测网络的收益。仿真案例验证了上述方法的有效性,并给出了各入侵检测系统的最优检测库配置方案。

本文只考虑了入侵检测系统之间的协作性,而攻击者是相互独立的,不存在任何合作。未来工作可以考虑更加智能的协作式攻击者,能通过分享信息来实时更新攻击策略,以及当任意一个攻击者攻击失败时,其余的攻击者可以协助攻击。

参考文献

[1] 邵子豪, 王慧强, 孟庆川, 等. 工业物联网安全态势评估方法研究综述[J]. 高技术通讯, 2020,30(9):908-917.

[2] YUN K, HUANG Q, MA Y X. Construction of network security perception system using Elman neural network [C]//2021 2nd International Conference on Computer Communication and Network Security. Xining, China: IEEE, 2021:187-190.

[3] 余正飞, 闫巧, 周鋆. 面向网络空间防御的对抗机器学习研究综述[J]. 自动化学报, 2022,48(7):1625-1649.

[4] 张博, 姚静, 梁旭辉. 入侵检测系统对计算机网络的安全维护[J]. 网络安全技术与应用, 2022(1):17-19.

[5] ZHANG R Y, SONG Y, WANG X H. Network intrusion detection scheme based on IPSO-SVM algorithm[C]// 2022 3rd Asia-Pacific Conference on Image Processing, Electronics and Computers. Dalian, China: IEEE, 2022: 1011-1014.

[6] HU S, XIONG Z Q. Research on multi-pattern matching algorithm of computer network intrusion detection system [C]//2021 International Conference on Computer Information Science and Artifical Intelligence. Kunming, China: IEEE, 2021:717-720.

[7] 马琳, 张莎莎, 宋姝雨, 等. 基于SDN的智能入侵检测系统模型与算法[J]. 高技术通讯, 2020,30(5):533-537.

[8] BAUSO D. Game theory: models, numerical methods and applications [M]. New York: Now Foundations and Trends, 2014:98-107.

[9] 赵晨馨, 董红召, 管宇辉, 等. 车辆网环境下车辆博弈换道协作策略[J]. 高技术通讯, 2021,31(1):64-74.

[10] WEILL C, OLIVEREAU A, ZEGHLACHE D. Configuration of the detection function in a distributed IDS using game theory[C]//2020 23rd Conference on Innovation in Clouds, Internet and Networks. Paris, France: IEEE, 2020:210-215.

[11] BENADDI H, IBRAHIMI K, BENSLIMANE A, et al. Robust enhancement of intrusion detection systems using deep reinforcement learning and stochastic game [J]. IEEE Transactions on Vehicular Technology, 2022, 71(10):11089-11102.

[12] FENG Y, SHOU Y H. Intrusion detection system configuration under asymmetric information[C]//2021 American Control Conference. New Orleans, USA: IEEE, 2021:4637-4642.

[13] GOVINDARAJ L, SUNDAN B, THANGASAMY A,et al.

An intrusion detection and prevention system for DDoS attacks using a 2-player Bayesian game theoretic approach [C]// 2021 4th International Conference on Computing and Communications Technologies. Chennai, India: IEEE, 2021:319-324.

[14] JIN R C, HE X F, DAI H Y. Collaborative IDS configuration: a two-layer game-theoretic approach [J]. IEEE Transactions on Cognitive Communications and Networking, 2018,4(4):803-815.

[15] XIA Z Q, TAN J J, GU K, et al. Detection resource allocation scheme for two-layer cooperative IDSs in smart grids[J]. Journal of Parallel and Distributed Computing, 2021,147:236-247.

[16] XIA Z Q, TAN J J, GU K. IDS intelligent configuration scheme against advanced adaptive attacks [J]. IEEE Transactions on Network Science and Engineering, 2021, 8(2):995-1008.

[17] FENG Y, SHOU Y H, YUX T. Jamming on remote estimation over wireless links under faded uncertainty: a Stackelberg game approach [J]. IEEE Transactions on Circuits and Systems Ⅱ: Express Briefs, 2021,68(7): 2593-2597.

[18] DIDDIGI R B, KAMANCHI C,BHATNAGAR S. A generalized minimax Q-learning algorithm for two-player zero-sum stochastic games [J]. IEEE Transactions on Automatic Control, 2022,67(9):4816-4823.

[19] LU J K, WANG X Y, WANG D Y, et al. Parallel Monte Carlo tree search in perfect informationgame with chance [C]// 2016 Chinese Control and Design Conference. Yinchuan, China: IEEE, 2016:5050-5053.

[20] 高阳, 陈世福, 陆鑫. 强化学习研究综述[J]. 自动化学报, 2004(1):86-100.

[21] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Cambridge: MIT Press, 2018:1-13.

[22] 谢识予. 经济博弈论[M]. 第 4 版. 上海: 复旦大学出版社, 2017:53-54.

[23] BELLMAN R. Dynamic programming [J]. Science, 1996,153(3731):34-37.

Research on configuration of detection libraries of cooperative intrusion detection systems under incomplete information

SHI Yuelou, YANG Danjie, FENG Yu, LI Yongqiang

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023)

Abstract

This paper studies the configuration problem of cooperative intrusion detection systems (IDS) with information limitation in finite-time horizon. Aiming at the optimal configuration of the detection libraries when facing different types of attacks and the contradiction in the allocation process, a method of two-layer detection libraries allocation scheme is proposed. In the first layer, the decision-making process of each intrusion detection system and the corresponding attacker are studied. In the second layer, the contradiction of loading detection libraries is solved by a centralized resource allocation method based on sharing strategy. In addition, the problem of the first layer can be solved in two steps. Firstly, by virtue of the fact that the attacker is not aware of the state, which leads to the information asymmetry, a belief-based stochastic game is constructed. And the strategy solved by the backward recursion algorithm is a stationary Nash equilibrium (SNE) strategy of the belief-based stochastic game. Secondly, the optimal detection libraries allocation plan can be solved by a hybrid Markov decision process. The simulation results show that the proposed algorithm is effective in obtaining the optimal configuration of detection libraries when facing different types of attacks.

Key words: incomplete information game, Markov decision process, intrusion detection system (IDS), resource allocation, network security