DOI:10.7544/issn1000-1239.2017.20170433
54(10):2344-2355,2017

基于专家系统的高级持续性威胁云端检测博弈

胡 晴[1,2] 吕世超[1,2] 石志强[1,2] 孙利民[1,2] 肖 亮[3]
[1](中国科学院大学网络空间安全学院 北京 100049)
[2](物联网信息安全技术北京市重点实验室(中国科学院信息工程研究所) 北京 100093)
[3](厦门大学通信工程系 福建厦门 361005)
(huqing@iie.ac.cn)

Advanced Persistent Threats Detection Game with Expert System for Cloud

Hu Qing[1,2], Lü Shichao[1,2], Shi Zhiqiang[1,2], Sun Limin[1,2], and Xiao Liang[3]
[1](*School of Cyber Security, University of Chinese Academy of Sciences, Beijing* 100049)
[2](*Beijing Key Laboratory of IOT Information Security Technology (Institute of Information Engineering, Chinese Academy of Sciences), Beijing* 100093)
[3](*Department of Communication Engineering, Xiamen University, Xiamen, Fujian* 361005)

Abstract Cloud computing systems are under threaten of advanced persistent threats (APT). It is hard for an autonomous detector to discover APT attacks accurately. The expert system (ES) can help to reduce detection errors via double-checking suspicious behaviors. However, it takes an extended period of time for the ES to recheck, which may lead to a defense delay. Besides, the ES makes mistakes too. In this paper, we discuss the necessity of the ES to participate in APT detection and defense for a cloud computing system by game theory, based on the consideration of miss detection rates and false alarm rates of both the APT detector and the ES. The ES-based APT detection method is designed, and the ES-APT game between an APT attacker and a defender is formulated. We derive its Nash equilibrium and analyze how the ES enhances the security of the cloud computing system. Also, the dynamic game is studied, in case that the APT attack model is unknowable. We present a reinforcement learning scheme for the cloud computing system with ES to get the optimal strategy. Simulation results show that, with the knowledge of the ES, both the defender's utility and the cloud computing system's security are improved compared with benchmark schemes.

Key words advanced persistent threats (APT); cloud security; expert system (ES); game theory; reinforcement learning

摘 要 云计算系统是高级持续性威胁(advanced persistent threats, APT)的重要攻击目标.自动化的APT检测器很难准确发现APT攻击,用专家系统对可疑行为进行二次检测可以减少检测错误.但是专

收稿日期:2017-06-10;**修回日期**:2017-08-01
基金项目:国家重点研发计划项目(2016YFB0800202);国防基础科研计划项目(JCKY2016602B001);国家自然科学基金项目(U1636120,61671396);北京市科委科技计划专项项目(Z161100002616032);CCF启明星辰鸿雁基金项目(2016-010)
This work was supported by the National Key Research and Development Program of China (2016YFB0800202), the National Defense Basic Scientific Research Program of China (JCKY2016602B001), the National Natural Science Foundation of China (U1636120, 61671396), Beijing Municipal Science and Technology Commission Program (Z161100002616032), and the CCF-Venustech Hongyan Research Initiative (2016-010).
通信作者:石志强(shizhiqiang@iie.ac.cn)

家系统完成二次检测需要花费一段额外的时间,可能导致防御响应延迟,而且专家系统本身也会产生误判.在综合考虑 APT 检测器和专家系统的虚警率和漏报率的基础上,用博弈论方法讨论在云计算系统的 APT 检测和防御中,利用专家系统进行二次检测的必要性.设计了一个基于专家系统的 APT 检测方案,并提出一个 ES-APT 检测博弈模型,推导其纳什均衡,据此研究了专家系统对云计算系统安全性能的改善作用.此外,当无法获得 APT 攻击模型时,提出了一种利用强化学习算法获取最优防御策略的方案.仿真结果表明:基于 WoLF-PHC 算法的动态 ES-APT 检测方案较之其他对照方案能够提高防御者的效用和云计算系统的安全性.

关键词　高级持续性威胁;云安全;专家系统;博弈论;强化学习

中图法分类号　TP393.08

随着云计算技术的发展,越来越多的数据被上传到云端,其中不乏金融、医疗、政务、通信、工业、农业等关系到国计民生的重要数据,导致云计算系统成为高级持续性威胁(advanced persistent threats,APT)的主要攻击目标.针对云计算系统的 APT 攻击主要是为了窃取机密信息.在达到目的之前,APT 攻击者会反复尝试,搜集大量目标系统的资料,并根据目标系统的防御情况不断调整攻击方案,直至成功[1].近年来,人们在 APT 防御方面做了大量研究.但实际情况表明,由于 APT 攻击不断尝试新的攻击手段、大量利用 0day 漏洞且擅于隐藏和擦除痕迹,很难准确检测到 APT 攻击.尤其是自动化的 APT 检测器,在工作过程中都会产生大量的虚警和漏报.虚警会导致错误的防御,给 APT 防御者带来人力、物力、财力以及时间上的损失.漏报更是为 APT 攻击继续深入提供便利,增加了攻击者窃密或实现其他攻击目的的机会.

为了缓解 APT 检测器的不准确性带来的危害,本文提出一种基于专家系统(expert system,ES)的 APT 攻击检测方案,简称 ES-APT 检测方案.专家系统一般是指计算机程序系统,用人工智能技术和计算机技术来模拟人类专家解决专业领域问题[2].本文的专家系统是由计算机专家系统和多个人类信息安全专家组成的多专家协作系统.在 ES-APT 检测方案中,APT 防御者借助 APT 检测器和专家系统对目标系统进行检测.APT 检测器持续扫描云计算系统,并根据防御者设置的时间间隔对所收集到的信息进行综合分析.当 APT 检测器报警时,触发专家系统进行二次检测,如果专家系统确认报警正确,则防御者采取措施阻断 APT 攻击,并修复由攻击造成的损失.从 APT 检测器报警到专家系统给出判断所经历的时间称为响应时间.在实际运行中,专家系统的判断也可能出错.

针对已知攻击模型的 APT 攻击,本文根据 ES-APT 检测方案提出一种 ES-APT 检测博弈模型,以 APT 攻击者和云计算系统的防御者为博弈的参与方.在该模型中,APT 检测器和专家系统的虚警率和漏报率是公共知识.APT 攻击者的策略是选择发动攻击的时机,防御者的策略是设置 APT 检测器进行综合分析的时间间隔.求解该模型的纳什均衡,可以得到防御者的最优策略.

动态的 ES-APT 检测博弈则用来研究无法获知 APT 攻击的攻击模型时 APT 防御者如何进行防御决策.本文提出一种基于 WoLF-PHC 算法的防御策略优化方案,并用模拟仿真验证了该方案的可行性和提升 APT 防御者效用的能力.

本文的主要贡献有 3 个方面:

1) 提出了一种 ES-APT 检测方案来缓解 APT 检测器的不准确性带来的危害,并基于该方案构建了一个以 APT 攻击者和防御者为参与人的 ES-APT 检测博弈模型;

2) 推导了 ES-APT 检测博弈的纳什均衡,并用数值分析揭示了 APT 检测器和专家系统的虚警率、漏报率,以及专家系统二次检测造成的防御延迟对博弈双方效用和云计算系统安全性的影响;

3) 在动态博弈中,基于 WoLF-PHC 算法设计了一种防御策略优化方案,用模拟仿真验证了该方案的可行性,并对比了该方案和其他对照方案的性能.

1　相关工作

博弈论在网络与信息安全相关领域应用广泛,涉及主动防御[3]、安全协议[4]、隐私保护[5-6]和攻击检测[7]等.在 APT 检测与防御方面,大量工作表明:博弈论是一种研究和解决 APT 攻击问题的有效方法.

文献[8]提出了一种防御隐蔽攻击的重复博弈

框架 FlipIt,研究了针对不同攻击策略的占优防御策略;文献[9]基于 FlipIt 框架研究了当 APT 攻防双方的时间、成本等资源受限时的近似最优防御策略,还提出了一个以防御者为主导者、攻击者为追随者的序贯博弈模型,设计了基于动态规划来获取防御者的近似最优策略的算法;文献[10]考虑了隐蔽攻击者逐步获取资源而防御者只能部分消除攻击立足点,且无法弥补任何已经发生的信息泄漏的情形,并构建博弈模型推导出最佳防御策略;文献[11]和文献[12]用前景理论论述了当 APT 攻防双方并非完全理性时,他们的主观程度对双方决策和效用的影响,设计了基于 Q-learning 的动态防御方案;文献[13]进一步用累积前景理论对 APT 攻防博弈进行了讨论;文献[14]分析了内部泄密者和 APT 攻击者的联合威胁,给出了可能存在内部泄密者时防御者的最优策略;文献[15]通过双层博弈模型研究攻击者与泄密者之间的交易以及攻击者与防御者之间的博弈,并求解了子博弈完美均衡;文献[16]用演化博弈论来捕捉长期连续的 APT 攻击行为,通过建立 2 个离散策略的 APT 防御博弈模型,研究了攻击策略和防御策略的动态稳定性.

然而,以上研究均未涉及检测 APT 攻击时可能出现的虚警和漏报. 实际应用中,在忽略 APT 检测的不准确性[17]的情况下做出的防御决策,可能会对防御效能产生负面影响. 本文提出 ES-APT 检测方案来提升 APT 检测的性能,并基于此构建 APT 攻击者和无法准确检测到攻击的防御者之间的博弈模型,从静态和动态 2 个方面为防御者提供更好的防御策略.

2 系统模型

本节介绍 ES-APT 检测方案以及基于此方案的 ES-APT 检测博弈的基本模型,并建立 APT 攻击者和防御者的效用函数.

2.1 ES-APT 检测方案

ES-APT 检测方案如图 1 所示. APT 检测器持续监听云计算系统的各类信息,并按防御者设定的检测时间间隔对这段时间内所监测到的数据进行综合分析,判断云计算系统是否已被攻击. 如果检测器认为系统没有遭受攻击,则防御者开始部署下一次检测时间间隔;反之,检测器给出告警,同时触发专家系统. 专家系统综合考量检测器收集的信息和其他与云计算系统相关的信息,进一步辨别系统是否安全. 只有专家系统确认了攻击确实发生,防御者才会采取防御措施对 APT 攻击进行阻断.

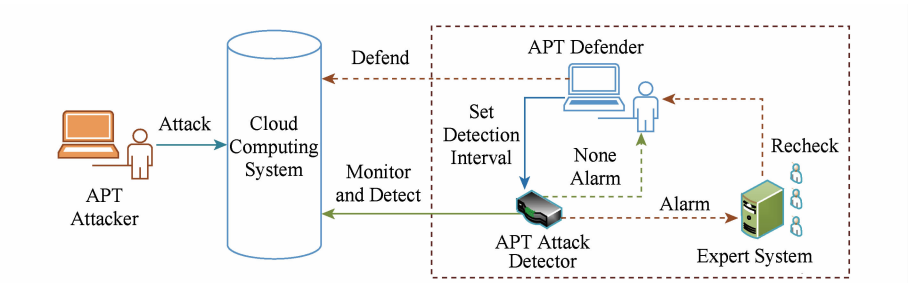


Fig. 1 The scheme of detecting APT attacks with an expert system

图 1 ES-APT 检测方案

2.2 基本模型

ES-APT 检测博弈是一个非合作博弈,有 2 个参与人:1)手段高明、隐蔽性强的 APT 攻击者;2)基于 ES-APT 检测方案进行防御的 APT 防御者. 假设在一次博弈的起始点,云计算系统处于安全状态. 攻击者和防御者基于对 APT 检测器和专家系统的虚警率、漏报率的考虑,在不知道对方如何决策的情况下,分别选择攻击时间 y 和检测时间间隔 x. 攻击者可以选择 $y=0$,即立刻攻击,而防御者不能选择 $x=0$,因为 APT 检测器根据 0 时间内的信息不可能判断是否存在攻击. 归一化之后有 $y\in[0,1]$, $x\in(0,1]$. 不论攻击者采用何种手段进行攻击,从其发动攻击到攻击生效都需要经历一段时间 z,且 $z>0$. 假设 APT 检测器和专家系统只能发现已经生效的攻击,其中 APT 检测器在检测时耗费的时间可以忽略不计,专家系统用于二次检测的耗时记为 t. ES-APT 检测博弈中部分可能出现的攻防互动情况如图 2 所示.

虚警是指系统未遭受攻击时被认为受到攻击,漏报则是系统遭受攻击后依然被认为处于安全状态. 若用 S 表示系统的真实状态,s 表示 APT 检测器判定的系统状态,s' 表示专家系统复检之后给出

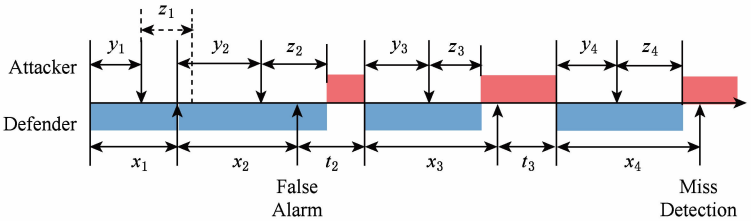


Fig. 2　Illustration of an ES-APT detection game

图 2　ES-APT 检测博弈示意图

的系统状态,下角标 0 和 1 分别指代未受攻击和受到攻击,则 APT 检测器的虚警率 p_m 和漏报率 p_f 分别为

$$p_m = Pr(s_0 \mid S_1), \tag{1}$$

$$p_f = Pr(s_1 \mid S_0). \tag{2}$$

专家系统的虚警率 p'_m 和漏报率 p'_f 分别为

$$p'_m = Pr(s'_0 \mid S_1), \tag{3}$$

$$p'_f = Pr(s'_1 \mid S_0), \tag{4}$$

以上 $Pr(\cdot\mid\cdot)$为条件概率.

APT 攻击者和防御者所争夺的云计算系统具有一定的价值,记为 C,其大小取决于该系统对攻击者和防御者而言的重要性. C 本为防御者所有,如果博弈的结局是云计算系统处于受攻击状态,则攻击者从防御者处夺走这部分价值.

2.3　效用函数

在推导效用函数之前,先给出一个度量 ES-APT 检测博弈性能的指标——安全率的定义.

定义 1. 安全率. 一次博弈中,云计算系统处于未受攻击状态的时间在博弈总时长中所占的比率称为安全率,记为 R.

在 ES-APT 检测博弈中,APT 防御者的效用由 4 部分组成:

1) 安全率带来的收益;

2) 从设定的检测时间间隔获益,间隔越长,APT 检测器收集的信息越多,越有利于 APT 检测器和专家系统做出正确判断,G_D 表示单位时间的获益;

3) 修复云计算系统所需的开销 C_R;

4) 如果博弈的最后已生效的 APT 攻击没有被发现,防御者输掉云计算系统价值 C.

攻击者的效用由 3 部分组成:

1) 安全率带来的损失;

2) 发动攻击时要付出的攻击成本 C_A;

3) 如果博弈的最后 APT 攻击生效且没有被阻断,攻击者获得云计算系统价值 C.

为了确定 ES-APT 检测博弈中防御方的效用函数 u_D 和攻击方的效用函数 u_A,我们将所有参数进行归一化处理,并分类讨论博弈中所有可能出现的情况. 从 APT 检测器准确性的角度,所有情况可归为四大类:检测器正确判定系统未受攻击、错误判定系统受到攻击、正确判定系统受到攻击和错误判定系统未受攻击.

1) 检测器正确判定系统未受攻击

如图 2 中序号为 1(即字母下角标为 1)的博弈所示,该情况出现的前提条件是 $y+z>x$,即在检测器检测之前,攻击尚未生效,其出现的概率是 $1-p_f$. 此时云计算系统的安全率为 1,防御者不需要进行修复操作,且不会失去 C. 这种情况下博弈双方的效用分别为

$$u_{D1}(x,y) = 1 + xG_D, \tag{5}$$

$$u_{A1}(x,y) = -1 - I(y \leqslant x)C_A, \tag{6}$$

其中,$I(\cdot)$为指示函数,括号内条件为真时 $I(\cdot)=1$,否则 $I(\cdot)=0$.

2) 检测器错误判定系统受到攻击

图 2 中序号为 2 的博弈是检测器错误判定系统受到攻击时,攻防双方可能的交互情况之一. 检测器错误判定系统受攻击的前提条件是 $y+z>x$,概率为 p_f. 检测器告警后,专家系统进行复验. 考虑到专家系统复验耗时较长,在其完成验证之前,原本没有生效的 APT 攻击可能会生效,所以云计算系统的安全率为 $\min((y+z)/(x+t))$.

如果 APT 攻击在专家系统复验完成前生效,即 $y+z \leqslant x+t$,专家系统的正确结论应该是云计算系统受到攻击. 专家系统判断正确的概率为 $1-p'_m$,出错的概率为 p'_m. 如果 APT 攻击一直没有生效,即 $y+z>x+t$,则专家系统判定云计算系统未受攻击的概率是 $1-p'_f$,判定受攻击的概率是 p'_f. 只要专家系统认为云计算系统受到攻击,防御者就会采取修复措施. 若专家系统漏报了攻击,则云计算系统将一直处于被攻击状态;若专家系统虚报了攻击,则可能会阻断已经发起但尚未生效的攻击. 所以,第 2 种情况下攻防双方的效用是

$$u_{D2}(x,y)=\min\left(\frac{y+z}{x+t},1\right)+xG_D-I(y+z\leqslant x+t)[p'_mC+(1-p'_m)C_R]-I(y+z>x+t)p'_fC_R, \quad (7)$$

$$u_{A2}(x,y)=-\min\left(\frac{y+z}{x+t},1\right)-I(y\leqslant x+t)C_A+I(y+z\leqslant x+t)p'_mC. \quad (8)$$

3) 检测器正确判定系统受到攻击

如图 2 中序号为 3 的博弈所示,检测器正确判定系统受到攻击的前提条件是 $y+z\leqslant x$,即在检测器检测之前,攻击已经生效,其出现的概率是 $1-p_m$. 此时检测器会触发专家系统进行验证,考虑到专家系统的响应时间,云计算系统的安全率为 $(y+z)/(x+t)$. 专家系统认同检测器的可能性是 $1-p'_m$,否定的可能性是 p'_m. 如果攻击被确认,防御者将修复云计算系统;反之,云计算系统得不到修复,其价值被攻击者夺走. 该情况下防御者和攻击者的效用分别为

$$u_{D3}(x,y)=\frac{y+z}{x+t}+xG_D-p'_mC-(1-p'_m)C_R, \quad (9)$$

$$u_{A3}(x,y)=-\frac{y+z}{x+t}-C_A+p'_mC. \quad (10)$$

4) 检测器错误判定系统未受攻击

图 2 中序号为 4 的博弈展现的是检测器错误判定系统未受攻击的情况,其前提条件是 $y+z\leqslant x$,概率为 p_m. 此时 APT 攻击被 APT 检测器漏掉,云计算系统被攻击者控制,攻防双方的效用为

$$u_{D4}(x,y)=\frac{y+z}{x}+xG_D-C, \quad (11)$$

$$u_{A4}(x,y)=-\frac{y+z}{x}-C_A+C. \quad (12)$$

综合以上分析可知,防御者的效用函数为

$$u_D(x,y)=I(y+z>x)[(1-p_f)u_{D1}+p_fu_{D2}]+I(y+z\leqslant x)[(1-p_m)u_{D3}+p_mu_{D4}], \quad (13)$$

攻击者的效用函数为

$$u_A(x,y)=I(y+z>x)[(1-p_f)u_{A1}+p_fu_{A2}]+I(y+z\leqslant x)[(1-p_m)u_{A3}+p_mu_{A4}]. \quad (14)$$

将式(5)(7)(9)(11)代入式(13),并整理可得:

$$u_D(x,y)=xG_D+I(y+z>x)\Big\{1-p_f+p_f\Big[\min\left(\frac{y+z}{x+t},1\right)-I(y+z\leqslant x+t)(p'_mC+(1-p'_m)C_R)-I(y+z>x+t)p'_fC_R\Big]\Big\}+I(y+z\leqslant x)\Big\{(1-p_m)\Big[\frac{y+z}{x+t}-p'_mC-(1-p'_m)C_R\Big]+p_m\left(\frac{y+z}{x}-C\right)\Big\}. \quad (15)$$

同样地,将式(6)(8)(10)(12)代入式(14),并整理可以得到攻击者的效用函数如下:

$$u_A(x,y)=I(y+z>x)\Big\{(1-p_f)[-1-I(y\leqslant x)C_A]+p_f\Big[-\min\left(\frac{y+z}{x+t},1\right)-I(y\leqslant x+t)C_A+I(y+z\leqslant x+t)p'_mC\Big]\Big\}+I(y+z\leqslant x)\Big[(1-p_m)\left(-\frac{y+z}{x+t}+p'_mC\right)-p_m\left(\frac{y+z}{x}-C\right)-C_A\Big]. \quad (16)$$

类似地,还可以得到安全率的表达式,如式(17)所示:

$$R(x,y)=I(y+z>x)\left[1-p_f+p_f\min\left(\frac{y+z}{x+t},1\right)\right]+I(y+z\leqslant x)\left[(1-p_m)\frac{y+z}{x+t}+p_m\frac{y+z}{x}\right]. \quad (17)$$

特别地,当 $t=0$, $p'_m=0$, $p'_f=1$ 时,意味着专家系统完全接受 APT 检测器的检测结果. 也就是说,这种情况下专家系统的作用仅是评估 APT 检测器的虚警率和漏报率,而不参与告警的二次检测.

3 混合策略 ES-APT 检测博弈

混合策略博弈是纯策略博弈的扩展. 运用混合策略可以增加博弈双方行为的不确定性,增加对方准确预测己方行动的难度. 本节详细介绍混合策略 ES-APT 检测博弈中攻防双方的策略空间,求解混合策略均衡,并通过数值分析研究混合策略下 ES-APT 检测方案的可行性和博弈的性能.

在混合策略 ES-APT 检测博弈中,APT 防御者从策略空间 $\{m/M\}_{1\leqslant m\leqslant M}$ 中选择检测时间间隔 x,APT 攻击者从策略空间 $\{n/N\}_{0\leqslant n\leqslant N}$ 中选择攻击时间间隔 y. 混合策略是指攻防双方各自按照一定概率,随机地从策略空间中选择一种纯策略作为实际的行动[18]. 因此,防御者的混合策略为 $\boldsymbol{\alpha}=[\alpha_m]_{1\leqslant m\leqslant M}$,其中 $\alpha_m=Pr(x=m/M)$ 是将 APT 检测时间间隔设为 x 的概率;攻击者的混合策略为 $\boldsymbol{\beta}=[\beta_n]_{0\leqslant n\leqslant N}$,其中 $\beta_n=Pr(y=n/N)$ 是将攻击时间间隔设为 y 的概率. 由混合策略的定义知:$\alpha_m\geqslant 0$, $\beta_n\geqslant 0$, $\sum_{m=1}^{M}\alpha_m=1$, $\sum_{n=0}^{N}\beta_n=1$.

一般而言,不论防御者还是攻击者都无法准确估算 APT 攻击发起之后,需要多长时间生效,亦即 z 是一个随机值. 为简便起见,以下把 z 看作常数.

混合策略博弈中的效用函数为期望效用函数.通过对式(15)和式(16)应用期望效用函数理论,得到防御者与攻击者的期望效用函数分别为

$$U_{\mathrm{D}}^{\mathrm{ETU}}(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{m=1}^{M}\sum_{n=0}^{N}\alpha_m\beta_n\times u_{\mathrm{D}}\left(\frac{m}{M},\frac{n}{N}\right),\tag{18}$$

$$U_{\mathrm{A}}^{\mathrm{ETU}}(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{m=1}^{M}\sum_{n=0}^{N}\alpha_m\beta_n\times u_{\mathrm{A}}\left(\frac{m}{M},\frac{n}{N}\right).\tag{19}$$

3.1　混合策略纳什均衡

用$(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$表示混合策略 ES-APT 检测博弈的纳什均衡,有:

$$\begin{cases}\boldsymbol{\alpha}^*=\arg\max\limits_{\boldsymbol{\alpha}}U_{\mathrm{D}}^{\mathrm{ETU}}(\boldsymbol{\alpha},\boldsymbol{\beta}^*),\\ \boldsymbol{\beta}^*=\arg\max\limits_{\boldsymbol{\beta}}U_{\mathrm{A}}^{\mathrm{ETU}}(\boldsymbol{\alpha}^*,\boldsymbol{\beta}),\\ \sum\limits_{m=1}^{M}\alpha_m=1,\boldsymbol{\alpha}\geqslant 0,\\ \sum\limits_{n=0}^{N}\beta_n=1,\boldsymbol{\beta}\geqslant 0.\end{cases}\tag{20}$$

定理 1. 如果式(21)的解存在,则式(21)中$(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$是混合策略 ES-APT 检测博弈的纳什均衡:

$$\begin{cases}\left[u_{\mathrm{D}}\left(\frac{m}{M},\frac{n}{N}\right)\right]\left[\beta_i^*\right]_{0\leqslant i\leqslant N}^{\mathrm{T}}=\lambda_{\mathrm{D}}\mathbf{1}_{N+1},\\ \left[u_{\mathrm{A}}\left(\frac{m}{M},\frac{n}{N}\right)\right]^{\mathrm{T}}\left[\alpha_k^*\right]_{1\leqslant k\leqslant M}^{\mathrm{T}}=\lambda_{\mathrm{A}}\mathbf{1}_{M},\\ \sum\limits_{m=1}^{M}\alpha_m=1,\boldsymbol{\alpha}\geqslant 0,\\ \sum\limits_{n=0}^{N}\beta_n=1,\boldsymbol{\beta}\geqslant 0,\\ \lambda_{\mathrm{D}}\geqslant 0,\lambda_{\mathrm{A}}\leqslant 0,\end{cases}\tag{21}$$

其中,$1\leqslant m\leqslant M,0\leqslant n\leqslant N,\mathbf{1}_\zeta$是一个$\zeta$维的元素全为 1 的列向量.

证明. 式(20)是一个有约束条件的优化问题,其拉格朗日函数L_{D}表示为

$$L_{\mathrm{D}}=U_{\mathrm{D}}^{\mathrm{ETU}}(\boldsymbol{\alpha},\boldsymbol{\beta}^*)-\phi\Big(\sum_{m=1}^{M}\alpha_m-1\Big)+\sum_{m=1}^{M}\mu_m\alpha_m.\tag{22}$$

其卡罗什-库恩-塔克(Karush-Kuhn-Tucker,KKT)条件为

$$\begin{cases}\dfrac{\partial L_{\mathrm{D}}}{\partial\alpha_m}=0,\\ -\alpha_m\leqslant 0,\mu_m\geqslant 0,\mu_m\alpha_m=0,1\leqslant m\leqslant M,\\ \sum\limits_{m=1}^{M}\alpha_m-1=0.\end{cases}\tag{23}$$

将式(23)与式(18)和式(22)联立可得:

$$\begin{cases}\sum\limits_{n=0}^{N}u_{\mathrm{D}}\left(\frac{i}{M},\frac{n}{N}\right)\beta_n^*=\lambda_{\mathrm{D}},i\in[1,M],\\ \sum\limits_{m=1}^{M}\alpha_m=1,\\ \lambda_{\mathrm{D}}\geqslant 0.\end{cases}\tag{24}$$

求解式(24)即可得到式(21)中的第 1 行.类似地,运用KKT条件可求得式(21)中的第2行.证毕.

为了使以上结论更为直观,我们在引理 1 中讨论了 ES-APT 检测博弈混合策略均衡的一个简单实例.

引理 1. $M=2,N=1$时,当且仅当条件I_1和I_2都成立时,式(25)和式(26)给出的$(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$是混合策略 ES-APT 检测博弈的唯一纳什均衡.

$$\alpha_1^*=\frac{u_{\mathrm{A}}(1,1)-u_{\mathrm{A}}(1,0)}{u_{\mathrm{A}}(\frac{1}{2},0)-u_{\mathrm{A}}(\frac{1}{2},1)+u_{\mathrm{A}}(1,1)-u_{\mathrm{A}}(1,0)},\tag{25}$$

$$\beta_0^*=\frac{u_{\mathrm{D}}(1,1)-u_{\mathrm{D}}(\frac{1}{2},1)}{u_{\mathrm{D}}(\frac{1}{2},0)-u_{\mathrm{D}}(1,0)+u_{\mathrm{D}}(1,1)-u_{\mathrm{D}}(\frac{1}{2},1)},\tag{26}$$

条件是:

$$I_1:\frac{u_{\mathrm{D}}(1,1)-u_{\mathrm{D}}(\frac{1}{2},1)}{u_{\mathrm{D}}(\frac{1}{2},0)-u_{\mathrm{D}}(1,0)}\geqslant 0$$

或:

$$\frac{u_{\mathrm{D}}(\frac{1}{2},0)-u_{\mathrm{D}}(1,0)}{u_{\mathrm{D}}(1,1)-u_{\mathrm{D}}(\frac{1}{2},1)}\geqslant 0,\tag{27}$$

$$I_2:\frac{u_{\mathrm{A}}(1,1)-u_{\mathrm{A}}(1,0)}{u_{\mathrm{A}}(\frac{1}{2},0)-u_{\mathrm{A}}(\frac{1}{2},1)}\geqslant 0$$

或:

$$\frac{u_{\mathrm{A}}(\frac{1}{2},0)-u_{\mathrm{A}}(\frac{1}{2},1)}{u_{\mathrm{A}}(1,1)-u_{\mathrm{A}}(1,0)}\geqslant 0.\tag{28}$$

将式(15)(16)代入式(25)(26)求解知,当$M=2,N=1$时,混合策略 ES-APT 检测博弈有唯一纳什均衡,由式(29)给出:

$$(\alpha_1^*,\beta_0^*)=\left(\frac{A_1}{A_2},\frac{B_1}{B_2}\right),\tag{29}$$

其中:

$$A_1=p_{\mathrm{f}}\left[1-\min\left(\frac{1+z}{1+t},1\right)\right]+I(z\leqslant t)p_{\mathrm{f}}p'_{\mathrm{m}}C+p_{\mathrm{m}}\frac{zt}{1+t}+(p_{\mathrm{m}}p'_{\mathrm{m}}-p_{\mathrm{m}}-p'_{\mathrm{m}})C+\frac{z}{1+t}-1,$$

$$A_2=p_{\mathrm{m}}\frac{zt}{1+t}+p_{\mathrm{f}}\left[\min\left(\frac{2+2z}{1+2t},1\right)-\min\left(\frac{1+z}{1+t},1\right)\right]-C_{\mathrm{A}}+I\left(z\leqslant\frac{1}{2}\right)\times\left[I\left(t-\frac{1}{2}<z\leqslant t\right)p_{\mathrm{f}}p'_{\mathrm{m}}C-\frac{2t}{1+2t}p_{\mathrm{m}}-\frac{z}{(1+2t)(1+t)}\right]+I\left(z>\frac{1}{2}\right)\times$$

$$\left[p_{\mathrm{f}}\left(1-\min\left(\frac{2z}{1+2t},1\right)\right)-p_{\mathrm{m}}(1-p'_{\mathrm{m}})C-p'_{\mathrm{m}}C+\frac{z}{1+t}-1\right]+\left[I\left(\frac{1}{2}<z\leqslant t\right)+I\left(\frac{1}{2}<z\leqslant t+\frac{1}{2}\right)\right]p_{\mathrm{f}}p'_{\mathrm{m}}C+I\left(t\geqslant\frac{1}{2}\right)p_{\mathrm{f}}C_{\mathrm{A}},$$

$$B_1=p_{\mathrm{f}}\left[\min\left(\frac{1+z}{1+t},1\right)-\min\left(\frac{2+2z}{1+2t},1\right)\right]+\frac{1}{2}G_{\mathrm{D}}+I\left(t-\frac{1}{2}<z\leqslant t\right)p_{\mathrm{f}}[p'_{\mathrm{f}}C_{\mathrm{R}}-p'_{\mathrm{m}}C-(1-p'_{\mathrm{m}})C_{\mathrm{R}}],$$

$$B_2=p_{\mathrm{f}}\left[\min\left(\frac{1+z}{1+t},1\right)-\min\left(\frac{2+2z}{1+2t},1\right)\right]-p_{\mathrm{m}}(z-C)-(1-p_{\mathrm{m}})\left[\frac{z}{1+t}-p'_{\mathrm{m}}C-(1-p'_{\mathrm{m}})C_{\mathrm{R}}\right]+I\left(t-\frac{1}{2}<z\leqslant t\right)p_{\mathrm{f}}[p'_{\mathrm{f}}C_{\mathrm{R}}-p'_{\mathrm{m}}C-(1-p'_{\mathrm{m}})C_{\mathrm{R}}]+I\left(z\leqslant\frac{1}{2}\right)\left[p_{\mathrm{m}}(2z-C)+(1-p_{\mathrm{m}})\left(\frac{2z}{1+2t}-p'_{\mathrm{m}}C-(1-p'_{\mathrm{m}})C_{\mathrm{R}}\right)\right]+I\left(z>\frac{1}{2}\right)\left[1-p_{\mathrm{f}}+p_{\mathrm{f}}\min\left(\frac{2z}{1+2t},1\right)\right]-I\left(\frac{1}{2}<z\leqslant t+\frac{1}{2}\right)\times p_{\mathrm{f}}[p'_{\mathrm{m}}C-(1-p'_{\mathrm{m}})C_{\mathrm{R}}]-I\left(z>t+\frac{1}{2}\right)p_{\mathrm{f}}p'_{\mathrm{f}}C_{\mathrm{R}}.$$

3.2 数值分析

本节用数值分析对混合策略 ES-APT 检测博弈的性能进行研究，主要关注 3 个指标：APT 防御者的效用、APT 攻击者的效用和云计算系统的安全率. 首先研究专家系统不参与决策时，APT 检测器的虚警率、漏报率对以上 3 个指标的影响；然后分析专家系统参与决策时检测器虚警率、漏报率的影响；最后讨论专家系统的响应时间、虚警率和漏报率对以上指标的影响. 为了达到更好的分析效果，本文选取的基本参数是 $G_{\mathrm{D}}=0.24$，$C=0.25$，$C_{\mathrm{R}}=0.1$，$C_{\mathrm{A}}=0.82$.

图 3 显示了专家系统不参与检测时，APT 检测器的漏报率和虚警率对混合策略 ES-APT 检测博弈性能的影响. 如图 3(a)所示，采用混合策略时，APT 防御者的效用不受 APT 检测器的漏报率影响，但随检测器虚警率的增加而降低，如检测器虚警率从 0 增加到 1 时，防御者效用从 1.12 减少到 1.02. 图 3(b)表明 APT 攻击者的效用不受检测器虚警率影响，而漏报率的上升能让攻击者效用增加，如检测器的漏报率从 0 增加到 1 时，攻击者的效用增加 10.3%. APT 检测器的漏报率和虚警率对云计算系统安全率的影响如图 3(c)所示. 当漏报率降低、虚警率增加时，安全率降低，尤其当漏报率接近 0、虚警率接近 1 时，云计算系统的安全率急剧下降. 这是因为，对攻击者而言，虚警率越高，APT 攻击发动之后、生效之前，因检测器虚警而被防御者阻断的可能性越大. 为了尽可能多地窃取信息，APT 攻击者必须加快攻击速度，让攻击尽可能在被检测器正确发现之前生效，从而更长时间控制系统. 对防御者而言，漏报率接近于 0、虚警率接近于 1 意味着几乎每次检测时检测器都会告警. 为了减少虚警出现的

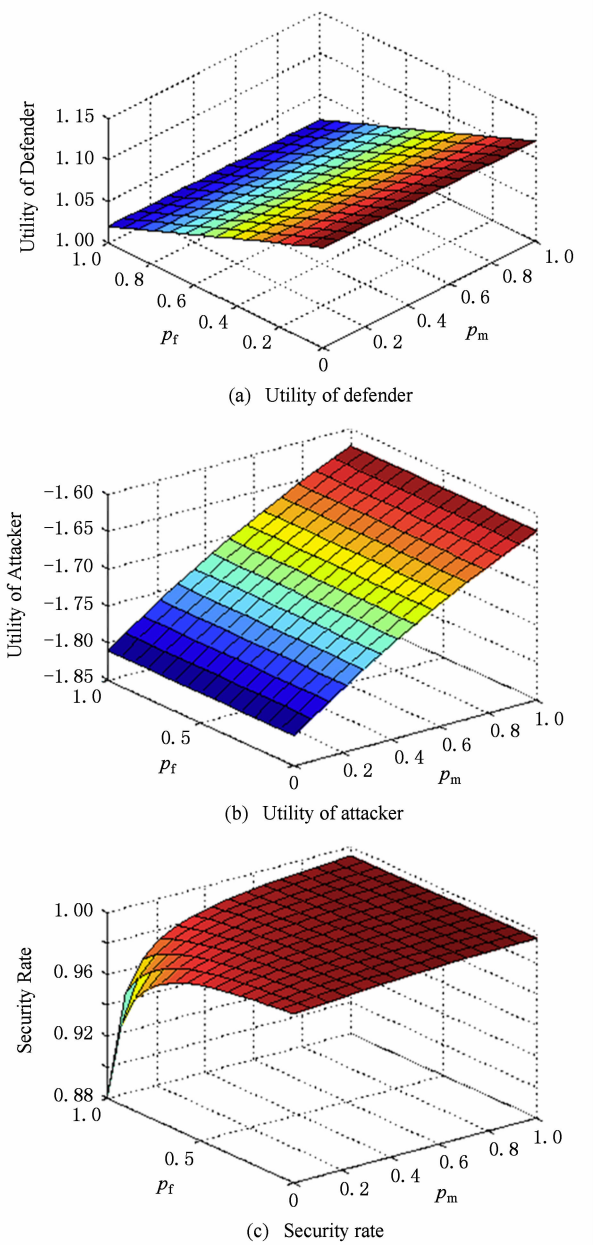


Fig. 3 Performance of the static game over error rates of the APT detector without ES

图 3 无专家系统时 APT 检测器错误率对静态博弈的影响

次数,防御者会延长检测周期.也就是说,在漏报率低虚警率高的情况下,攻击者会加快攻击速度,而防御者会延长防御周期,从而导致安全率急剧下降.

为了更清晰地说明专家系统参与 APT 检测后,对混合策略 ES-APT 检测博弈的性能造成了正面影响还是负面影响,我们计算了专家系统参与检测和不参与检测时博弈性能的差值.图 4 展示的即是该差值随 APT 检测器的漏报率和虚警率变化的情况.不失一般性,在图 4 中取 $t=0.1$,$p'_{\mathrm{m}}=0.05$,$p'_{\mathrm{f}}=0.1$.图 4(a)表明,防御者的效用之差始终为正值,不受 APT 检测器漏报率的影响,随检测器虚警率的增加而增加,如虚警率从 0 增加到 1 时,防御者效用之差从 0 增加到 0.09.说明专家系统可以提高防御者的效用,减缓由检测器虚警率导致的防御者效用的下降.图 4(b)表明,攻击者的效用之差也始终为正值,不受检测器虚警率的影响,但随着检测器漏报率的增加而降低,如漏报率从 0 增加到 1 时,攻击者效用之差从 0.03 减少到 0.说明虽然专家系统的引入会让攻击者获益,但专家系统能减少攻击者从检测器漏报率的增加中获得的利益.图 4(c)展示了云计算系统安全率之差的变化情况.安全率之差随着检测器虚警率的增加而增加,而与检测器的漏报率的关系比较复杂,不是单调递增或单调递减.比如,当 $p_{\mathrm{m}}=0$ 时,p_{f} 从 0 增加到 0.75 左右,安全率之差随之从 −0.02 增加到 0,而随着 p_{f} 继续增加直到等于 1,安全率之差最终上升为 0.08.说明当检测器虚警率较低时,专家系统参与检测反而会对云计算系统的安全率产生负面影响,而当检测器虚警率较高时专家系统对安全率的促进作用非常明显.此外,当检测器的虚警率不变时,随着其漏报率的增加,安全率先下降后上升,虚警率越大这种变化越明显.比如 $p_{\mathrm{f}}=1$ 时,随着 p_{m} 从 0 上升到 0.6,安全率的差值从 0.085 下降到几乎为 0;而在 p_{m} 继续上升到 1 的过程中,安全率的差值从接近于 0 上升到 0.002;当 p_{m} 取值在 0.367～0.700 之间时,安全率的差值为接近于 0 的负数,其余情况下为正数.将图 3(c)与图 4(c)结合起来看,可知专家系统能使安全率处于一个比较稳定的水平,减少检测器的不准确性对安全率的影响.

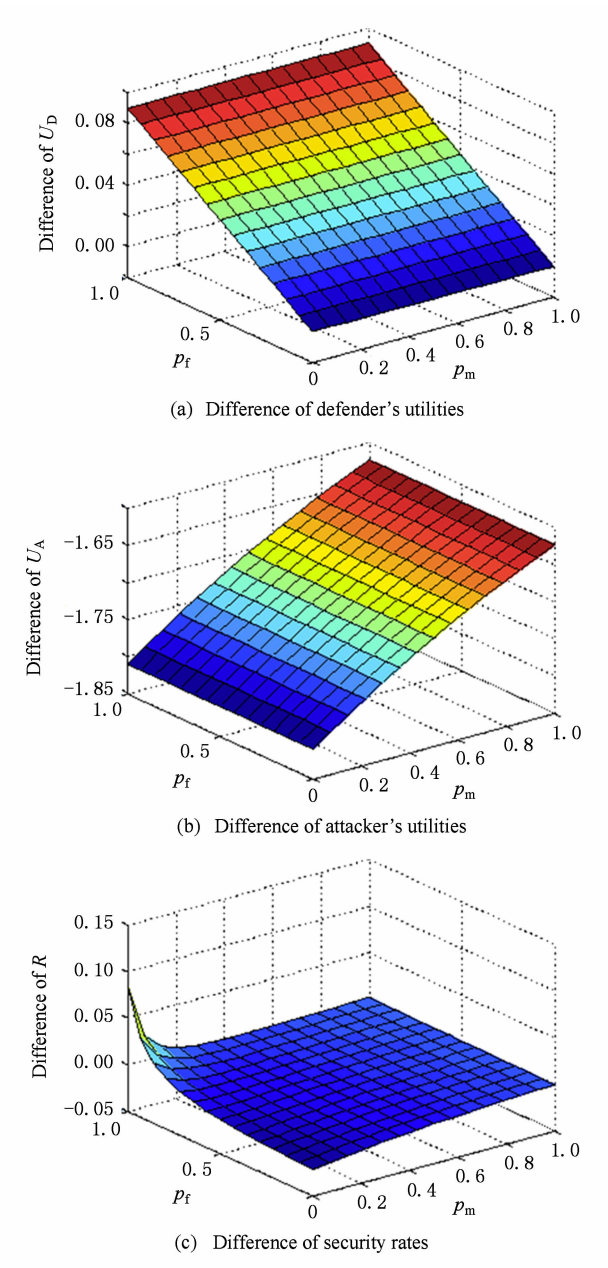


Fig. 4　Performance difference of the static game over error rates of the APT detector between with and without ES

图 4　有无专家系统情况下 APT 检测器错误率对静态博弈影响之差

专家系统的性能对攻防双方效用和安全率的影响如图 5 所示.总体来看,专家系统的响应时间不同,攻防双方效用和安全率受专家系统性能的影响也略有不同.图 5(a)显示了防御者效用受专家系统性能影响的情况.防御者的效用被响应时间 t 取值为 0.4 和 0.9 分为 3 段.$t=0.4$ 时防御者效用瞬间跳升,而 $t=0.9$ 时防御者效用瞬间下降.在每一个区间段中,防御者的效用都随响应时间的变长而降低.专家系统的漏报率和虚警率增加时,防御者效用降低.如 $t=0.3$ 时,随着 p'_{m} 从 0.1 上升到 0.7、p'_{f} 从 0.1 上升到 0.6,防御者效用降低了 4.1%.这说明专家系统性能越差,对防御者效用的负面作用越大.图 5(b)显示了攻击者效用受专家系统性能影响

的情况,其中也有 2 个跳跃点,t 取值为 0.5 和 0.9. 如图 5(b)所示:专家系统的响应时间越长,攻击者效用越大,如 t 从 0 上升到 0.5 时,攻击者效用上升大约 10.1%;攻击者效用不受专家系统的虚警率影响,但是随着其漏报率增加而增加,例如 $t=0.3$ 时,随着 p'_m 从 0.1 上升到 0.7,攻击者效用提升大约 5.8%. 图 5(c)展示了专家系统性能对云计算系统安全率的影响,该图中有 3 个跳跃点,当 $t=0.4$ 和 $t=0.5$ 时安全率突然上升,$t=0.9$ 时安全率突然下降. 在这 3 个点分割成的 4 个区间段内,安全率均随响应时间的增长而下降,整体上安全率也是随 t 的增大而变小. 虽然在 $t=0.3$ 到 $t=0.6$ 之间,安全率并不是随着 t 单调递减,但这个区间内的安全率最大值和最小值之间差别并不明显,2 次跳跃上升对改善性能意义不大. 因此,要维持较高的安全率,最重要的是将响应时间控制在 0.3 以内. 此外,从图 5(c)还可以看出,随着专家系统漏报率的增加和虚警率的降低,云计算系统的安全率上升. 比如当 $t=0.3$ 时,随着 p'_m 从 0.1 上升到 0.7 以及 p'_f 从 0.6 下降到 0.1,安全率上升了 3.9%.

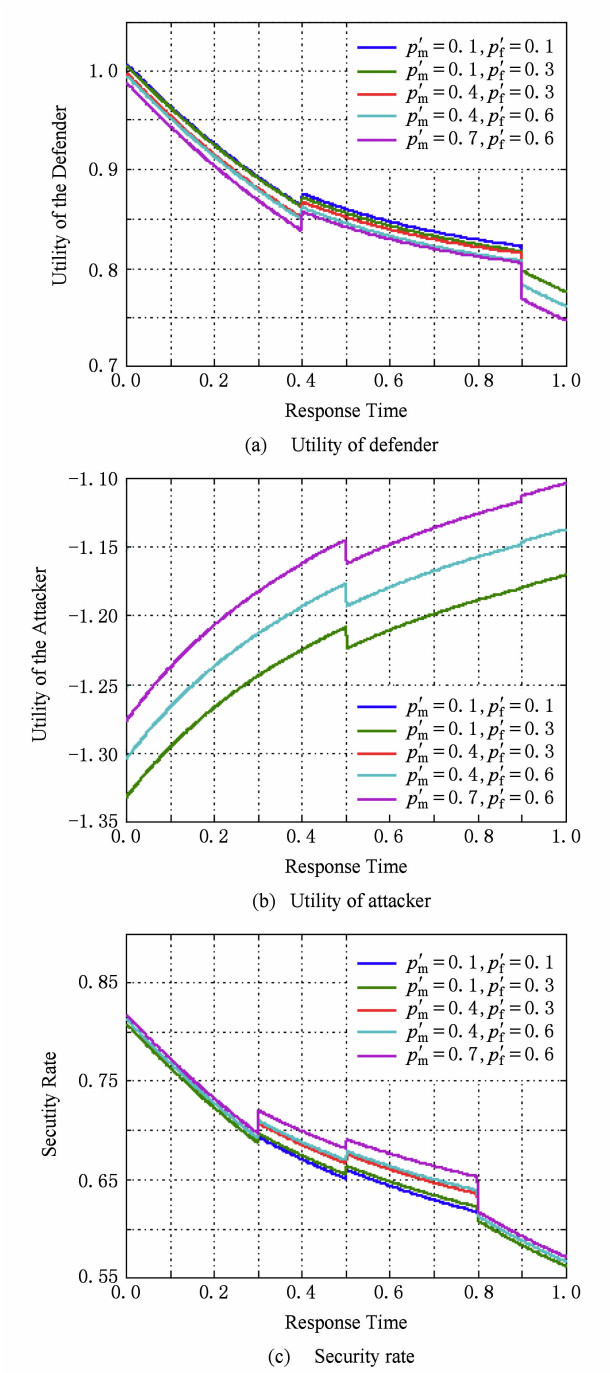


Fig. 5 Performance of the static game over the response time and error rates of ES

图 5 专家系统性能对静态博弈的影响

综上所述,引入专家系统进行二次检测,可以缓解 APT 检测器的虚警和漏报给防御者效用以及云计算系统安全率造成的负面影响,提升防御者效用并减少 APT 检测器的虚警和漏报造成的安全率的波动. 而为了使专家系统发挥更好的作用,必须提升专家系统的性能,减少响应时间,降低其漏报率和误报率. 因此,在与 APT 攻击者的对抗中,专家系统必须不断学习,扩充知识库,对 APT 攻击者的攻击手段进行深入研究,关注并预测新的攻击方法,尽可能先于攻击者发现 0day 漏洞等.

4 动态 ES-APT 检测博弈

APT 攻击者为了达到攻击目的会不断尝试新的方法. 因此,在实际中很多 APT 攻击者的攻击模型是未知的,其攻击成本、攻击生效时间等因素也不确定. 为了应对这种情况,我们用动态 ES-APT 检测博弈来分析攻击者与防御者之间的行为交互,提出一种基于强化学习算法,即赢或加速学习策略爬山算法(win or learn faster policy hill-climbing, WoLF-PHC)的最优决策方案. 在动态 ES-APT 检测博弈中,防御者用基于 WoLF-PHC 的最优决策方案来选择防御策略.

策略爬山(policy hill-climbing, PHC)算法是 Q-learning 算法的扩展,提升了其学习效率. 而 WoLF-PHC 则通过将赢或加速学习(win or learn faster, WoLF)原则用到 PHC 算法上,进一步提高了算法的收敛性[19]. WoLF-PHC 和 Q-learning 一样是离策略算法,不依赖系统模型,且都通过式(30)更新质量矩阵

$$\boldsymbol{Q}(s,x)\leftarrow(1-\mu)\boldsymbol{Q}(s,x)+\mu(u_{\mathrm{D}}+\gamma\max_{x'}\boldsymbol{Q}(s',x')),\tag{30}$$

其中,s 是状态,x 是防御者的动作,u_{D} 表示防御者的瞬时效用.在动态 ES-APT 检测博弈中,用攻击的整个周期表示系统状态,即 $s=y+z$.最大 $\boldsymbol{Q}$ 值通过 ε-greedy 算法选取,即:

$$Pr(x=\hat{x})=\begin{cases}1-\varepsilon, & \hat{x}=\arg\max_{x}\boldsymbol{Q}(s,x),\\ \dfrac{\varepsilon}{M-1}, & \text{otherwise.}\end{cases}\tag{31}$$

其中,$\varepsilon\in(0,1)$,通常是一个很小的正数,M 是防御者策略空间中动作的总个数.

基于 WoLF-PHC 的动态 ES-APT 检测方案见算法 1.

算法 1. 基于 WoLF-PHC 的动态 ES-APT 检测.

1) 初始化所有参数:$\mu=0.75$, $\gamma=0.7$, $\delta_{\mathrm{l}}=0.4$, $\delta_{\mathrm{w}}=0.2$, $\varepsilon=0.1$, $x\in\boldsymbol{\alpha}$, $y+z=0$, $t=0.1$, $\boldsymbol{Q}(s,x)\leftarrow0$, $\pi(s,x)\leftarrow\dfrac{1}{|\boldsymbol{\alpha}|}$, $C(s)\leftarrow0$;

2) for $k=1,2,3,\cdots$ do

3)　更新状态 s,$s=y+z$;

4)　对应 s,以概率 $\pi(s,x)$ 选择动作 x;

5)　依据 x 对云计算系统进行检测;

6)　观察 u_{D} 和接下来的状态 s,更新状态 s;

7)　依据式(30)更新 $\boldsymbol{Q}$;

8)　更新平均策略 $\bar{\pi}$,$C(s)\leftarrow C(s)+1$, $\forall x'\in\boldsymbol{\alpha}$, $\bar{\pi}(s,x')\leftarrow\bar{\pi}(s,x')+\dfrac{\pi(s,x')-\bar{\pi}(s,x')}{C(s)}$;

9)　通过 $\pi(s,x)\leftarrow\pi(s,x)+\Delta$ 更新 $\pi(s,x)$,

$$\Delta=\begin{cases}\delta, & x=\arg\max_{x'}\boldsymbol{Q}(s,x'),\\ \dfrac{-\delta}{|\boldsymbol{\alpha}|-1}, & \text{otherwise},\end{cases}$$

$$\delta=\begin{cases}\delta_{\mathrm{w}}, & \sum_{x}\pi(s,x)\boldsymbol{Q}(s,x)>\sum_{x}\bar{\pi}(s,x)\boldsymbol{Q}(s,x),\\ \delta_{\mathrm{l}}, & \text{otherwise};\end{cases}$$

10) end for

我们用基于 Q-learning 的动态 ES-APT 检测方案[20]作为对照,如算法 2 所示.

算法 2. 基于 Q-learning 的动态 ES-APT 检测.

1) 初始化所有参数:$\mu=0.75$, $\gamma=0.7$, $\varepsilon=0.1$, $y+z=0$, $\boldsymbol{Q}(s,x)\leftarrow0$;

2) for $k=1,2,3,\cdots$ do

3)　更新状态 s,$s=y+z$;

4)　通过式(31)选择动作 x;

5)　依据 x 对云计算系统进行检测;

6)　观察 u_{D} 和接下来的状态 s,更新状态 s;

7)　依据式(30)更新 $\boldsymbol{Q}$;

8) end for

5　模拟仿真

本节通过模拟仿真来评价动态 ES-APT 检测博弈中基于 WoLF-PHC 的动态 ES-APT 检测方案的性能,并且与基于 Q-learning 的检测方案和 ε-greedy 检测算法作对比.为了达到良好的性能,将 WoLF-PHC 算法的参数设置为 $\mu=0.75$,$\gamma=0.7$,$\delta_{\mathrm{l}}=0.4$,$\delta_{\mathrm{w}}=0.2$,$\varepsilon=0.1$;Q-learning 算法的参数设置为 $\mu=0.75$,$\gamma=0.7$,$\varepsilon=0.1$;ε-greedy 检测算法中 $\varepsilon=0.1$.检测和攻击参数设置为典型值,$p_{\mathrm{m}}=0.45$,$p_{\mathrm{f}}=0.1$,$p'_{\mathrm{m}}=0.15$,$p'_{\mathrm{f}}=0.6$,$G_{\mathrm{D}}=0.24$,$C=0.25$,$C_{\mathrm{R}}=0.1$,$C_{\mathrm{A}}=0.82$,$t=0.1$.在模拟仿真中,假设攻击者总是追求效用最大化,每次都依据上一次防御者实施攻击检测的时间 x 来选择能让自己的效用,如式(16)所示,最大的攻击时间间隔 y,即:

$$Pr(y=\hat{y})=\begin{cases}1-N\upsilon, & \hat{y}=\arg\max_{y}u_{\mathrm{A}}(\hat{x},y),\\ \upsilon, & \text{otherwise.}\end{cases}\tag{32}$$

仿真结果如图 6 所示.图 6(a)展示的是防御者的效用随实验方案运行次数的变化.基于 WoLF-PHC 动态检测方案,防御者的效用在 15 次之后收敛到 1.125 左右,400 次的平均效用约为 1.116.当采用 Q-learning 方法时,防御者的效用在 35 次之后收敛到 1.075 左右,400 次的平均效用约为 1.064.Q-learning 方法的平均效用比 WoLF-PHC 方法低大约 4.9%,收敛速度也明显较慢.基于 ε-greedy 算法,防御者的效用一直维持在 0.995 上下,其平均效用比 WoLF-PHC 低 10.8%.

从图 6(b)可知,当防御者基于 WoLF-PHC 部署动态的 ES-APT 检测方案时,云计算系统的安全率从 0.860 逐步上升到 1,在算法运行大约 18 次时收敛,整个 400 次运行过程中安全率的平均值为 0.994.基于 Q-learning 检测算法,安全率最终能与 WoLF-PHC 达到同样水平,400 次的平均值为 0.993,但是 Q-learning 算法收敛较慢,在大约 30 次左右收敛.而基于 ε-greedy 算法,安全率一开始就能上升到 0.90 左右,但最终也只能维持在这个水平,其 400 次的平均值比 WoLF-PHC 检测方案低约 10%.

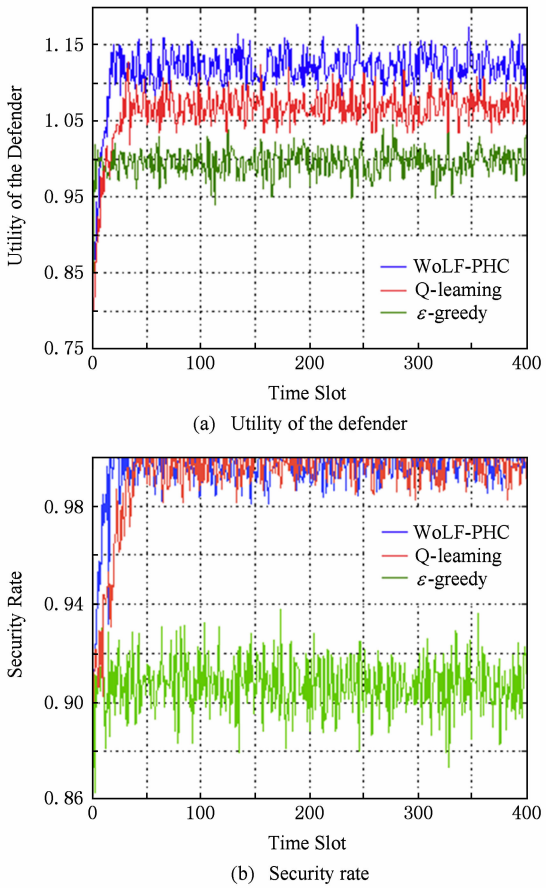


Fig. 6 Performance of the dynamic ES-APT detection game

图 6 动态 ES-APT 检测博弈性能图

从图 6 结果可以看出,基于 WoLF-PHC 的动态 ES-APT 检测方案比 Q-learning 的收敛性好,而且与 2 种对照方案相比,能明显提高防御者的效用和云计算系统的安全率.

6 总 结

本文提出了一种基于专家系统的 APT 检测方案,并在此基础上建立了 2 种 ES-APT 检测博弈,一个静态博弈和一个动态博弈,求解了静态博弈的混合策略均衡,并用数值分析研究了其性能. 数值分析结果显示,虽然专家系统的响应时间和虚警、漏报率对云计算系统的安全率以及攻击者的效用有一定的负面影响,但总体来说,基于专家系统的 APT 检测方案能够消除因 APT 检测器的不准确性造成的安全率和防御者效用的降低. 通过提升专家系统的性能,可以更好地改善云计算系统的安全性能. 在动态博弈中,基于 WoLF-PHC 算法设计了一种 ES-APT 动态检测方案,并与基于 Q-learning 和 ε-greedy 算法的方法进行了比较. 仿真结果表明:在 ES-APT 动态博弈中,基于 WoLF-PHC 的 ES-APT 动态检测方案能让防御者优化其策略,达到更好的防御效果. 与 Q-learning 相比,WoLF-PHC 能让防御者更快地获得其最优策略. 较之 Q-learning 和 ε-greedy,WoLF-PHC 能提高防御者的效用,同时也让云计算系统的安全率更高.

参 考 文 献

[1] Cole E. Advanced Persistent Threat: Understanding the Danger and How to Protect Your Organization [M]. Rockland, Massachusetts: Syngress Publishing, 2012: 11-36

[2] Coombs M J, Bolc L. Expert System Applications [M]. Berlin: Springer, 1988: 55-63

[3] Lin Wangqun, Wang Hui, Liu Jiahong, et al. Research on active defense technology in network security based on non-cooperative dynamic game theory [J]. Journal of Computer Research and Development, 2011, 48(2): 306-316 (in Chinese)
(林旺群, 王慧, 刘家红, 等. 基于非合作动态博弈的网络安全主动防御技术研究[J]. 计算机研究与发展, 2011, 48(2): 306-316)

[4] Tian Youliang, Peng Changgen, Ma Jianfeng, et al. Game-theoretic mechanism for cryptographic protocol [J]. Journal of Computer Research and Development, 2014, 51(2): 344-352 (in Chinese)
(田有亮, 彭长根, 马建峰, 等. 安全协议的博弈论机制[J]. 计算机研究与发展, 2014, 51(2): 344-352)

[5] He Yunhua, Sun Limin, Yang Weidong, et al. A game theory-based analysis of data privacy in vehicular sensor networks [J]. International Journal of Distributed Sensor Networks, 2014, 10(1): 1-14

[6] He Yunhua, Sun Limin, Yang Weidong, et al. Privacy preserving for node trajectory in VSN: A game-theoretic analysis based approach [J]. Journal of Computer Research and Development, 2014, 51(11): 2483-2492 (in Chinese)
(何云华, 孙利民, 杨卫东, 等. 基于博弈分析的车辆感知网络节点轨迹隐私保护机制[J]. 计算机研究与发展, 2014, 51(11): 2483-2492)

[7] Wang Yichuan, Ma Jianfeng, Lu Di, et al. Game optimization for internal DDoS attack detection in cloud computing [J]. Journal of Computer Research and Development, 2015, 52(8): 1873-1882 (in Chinese)
(王一川, 马建峰, 卢笛, 等. 面向云环境内部 DDoS 攻击检测的博弈论优化[J]. 计算机研究与发展, 2015, 52(8): 1873-1882)

[8] Marten V D, Ari J, Oprea A, et al. Flipit: The game of stealthy takeover [J]. Journal of Cryptology, 2013, 26(4): 655-713

[9] Zhang Ming, Zheng Zizhan, Shroff N B. A game theoretic model for defending against stealthy attacks with limited resources [C] //Proc of the 6th Decision and Game Theory for Security. Berlin: Springer, 2015: 93-112

[10] Farhang S, Grossklags J. Flipleakage: A game-theoretic approach to protect against stealthy attackers in the presence of information leakage [C] //Proc of the 7th Decision and Game Theory for Security. Berlin: Springer, 2016: 195-214

[11] Xu Dongjin, Li Yanda, Xiao Liang, et al. Prospect theoretic study of cloud storage defense against advanced persistent threats [C] //Proc of the 60th Global Communications Conf. Piscataway, NJ: IEEE, 2017: 1-6

[12] Xiao Liang, Xu Dongjin, Xie Caixia, et al. Cloud storage defense against advanced persistent threats: A prospect theoretic study [J]. IEEE Journal on Selected Areas in Communications, 2017, 35(3): 534-544

[13] Xu Dongjin, Xiao Liang, Mandayam N B, et al. Cumulative prospect theoretic study of a cloud storage defense game against advanced persistent threats [C] //Proc of the 36th IEEE Int Conf on Computer Communications (IEEE INFOCOM WKSHPS 2017). Piscataway, NJ: IEEE, 2017

[14] Hu Pengfei, Li Hongxing, Fu Hao, et al. Dynamic defense strategy against advanced persistent threat with insiders [C] //Proc of the 34th Int Conf on Computer Communications (IEEE INFOCOM 2015). Piscataway, NJ: IEEE, 2015: 747-755

[15] Feng Xiaotao, Zheng Zizhan, Hu Pengfei, et al. Stealthy attacks meets insider threats: A three-player game model [C] //Proc of the 34th Military Communications Conf (IEEE MILCOM 2015). Piscataway, NJ: IEEE 2015: 25-30

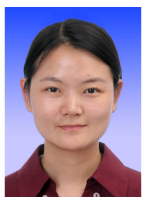
[16] Abass A, Xiao Liang, Mandayam N B, et al. Evolutionary game theoretic analysis of advanced persistent threats against cloud storage [J]. IEEE Access, 2017, 5(1): 8482-8491

[17] Xiao Liang, Li Yan, Han Guoan, et al. Phy-layer spoofing detection with reinforcement learning in wireless networks [J]. IEEE Trans on Vehicular Technology, 2016, 65(12): 10037-10047

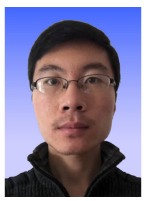
[18] Osborne M J, Rubinstein A. A Course in Game Theory [M]. Cambridge, Massachusetts: MIT Press, 1994: 29-40

[19] Bowling M, Veloso M. Rational and convergent learning in stochastic games [C] //Proc of the 33rd Int Joint Conf on Artificial Intelligence. San Francisco: Margan Kaufmann, 2001: 1021-1026

[20] Hu Qing, Lü Shichao, Shi Zhiqiang, et al. Defense against advanced persistent threats with expert system for Internet of things [G] //LNCS 10251: Proc of the 12th Int Conf on Wireless Algorithms, Systems, and Applications. Berlin: Springer, 2017: 326-337



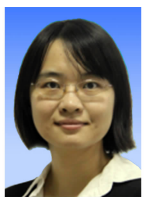
Hu Qing, born in 1985. PhD candidate. Member of CCF. Her main research interests include advanced persistent threats and IOT security.



Lü Shichao, born in 1985. PhD candidate, engineer. Member of CCF. His main research interests include wireless communication systems security (lvshichao@iie.ac.cn).

Shi Zhiqiang, born in 1970. PhD, senior engineer, PhD supervisor. Senior member of CCF. His main research interests include industrial control system security, cyber security, etc.

Sun LiMin, born in 1966. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include IOT security, cyber security, etc (sunlimin@iie.ac.cn).



Xiao Liang, born in 1980. PhD, professor, PhD supervisor. Senior member of CCF. Her main research interests include network security, wireless communications, smart grids, etc (Lxiao@xmu.edu.cn).