基于不完全信息动态博弈的动态目标防御最优策略选取研究

刘 江[1,2],张红旗[1,2],刘 艺[1,2]
(1. 解放军信息工程大学,河南郑州 450001;2. 河南省信息安全重点实验室,河南郑州 450001)

摘 要: 针对动态目标防御的最优策略选取问题,分析了动态目标防御环境下的攻防对抗特点,提出了动态目标防御策略的收益量化方法,基于不完全信息动态博弈构建了面向动态目标防御的单阶段和多阶段博弈模型,给出了精炼贝叶斯均衡求解算法和先验信念修正方法,获得了不同安全态势下的最优动态目标防御策略. 最后,通过实例说明和验证了上述模型和方法的可行性和有效性,总结了利用动态目标防御策略进行网络防御的一般性规律.

关键词: 动态目标防御; 策略选取; 不完全信息动态博弈; 精炼贝叶斯均衡
中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2018)01-0082-08
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2018.01.012

Research on Optimal Selection of Moving Target Defense Policy Based on Dynamic Game with Incomplete Information

LIU Jiang[1,2], ZHANG Hong-qi[1,2], LIU Yi[1,2]
(1. *PLA Information Science and Technology University, Zhengzhou, Henan* 450001, *China*;
2. *Henan Key Laboratory of Information Security, Zhengzhou, Henan* 450001, *China*)

Abstract: This paper is focused on the optimal policy selection for moving target defense. Attack-defense confrontation in moving target defense environment is analyzed. Reward quantization method of moving target defense policy is proposed. Single-stage and multi-stage moving target defense game models are constructed based on the dynamic game with incomplete information. The algorithm to obtain perfect Bayesian equilibrium and the method to revise prior belief are proposed. Optimal moving target defense policies are obtained under different security situations. Finally, not only the feasibility and effectiveness of the proposed model and method are illustrated and verified in a representative example but also general rules of network defense using moving target defense policies are summarized.

Key words: moving target defense; policy selection; dynamic game with incomplete information; perfect Bayesian equilibrium

1 引言

传统网络防御技术根据攻击特征配置和下发防御策略,属于被动滞后的防御手段,导致防御一方在网络空间安全对抗中长期处于劣势地位. 现有研究表明,信息系统属性的确定性、静态性和相似性是造成这一现状的重要原因[1]. 为扭转网络攻防地位的不对称性,美国科学技术委员会提出动态目标防御技术 MTD(Moving Target Defense)[2],其核心思想是通过构建、评价和部署持续变换的多样性机制和策略,增加信息系统的多样性、动态性和随机性,增大攻击难度和代价,有效限制系统漏洞持续暴露及被利用的机会,提高信息系统的抗攻击性[3].

与其它防御策略类似,MTD 在提升网络主动防御能力的同时,也不可避免地增加了防御成本,特别是周期性的 MTD 策略切换更会造成系统功能损耗的增加. 针对不同安全威胁和安全态势,实施具有合理跳变对象和跳变周期的 MTD 策略,以最大化防御收益成为

收稿日期:2016-07-01;修回日期:2017-03-29;责任编辑:李勇锋
基金项目:国家 863 高技术研究发展计划(No. 2012AA012704);郑州市科技领军人才基金(No. 131PLJRC644)

MTD 领域的研究热点. 此外,网络空间对抗所展现出的目标对立性、策略依存性和关系非合作性与博弈论的基本特征相一致[4]. 因此,本文拟采用博弈论描述和分析 MTD 防御环境下的安全状态推演,解决最优 MTD 防御策略选取问题.

文献[5]利用博弈论和机器学习分析了应对具有自学习能力攻击者的 MTD 策略,指出均匀随机化 MTD 策略是其最优解. 但是,其重点针对具有预测能力的攻击者,忽略了 MTD 策略自身的多样性和动态性. 文献[6]定义了攻击面转移的概念并对其进行了量化分析,将 MTD 防御环境下的攻防交互建模为一种单阶段完全信息静态博弈,将最优 MTD 策略视为系统安全性与可用性之间的一种均衡. 但是,其对攻防双方完全信息的假设与实际攻防对抗不符,且是一种单阶段博弈模型. 文献[7]提出基于反馈信息架构的多阶段防御机制,构建了二人零和博弈模型,重点分析了防御者改变攻击面和攻击者改变攻击向量的成本,但对攻防策略的收益量化分析不够全面,未能体现 MTD 策略特点. 文献[8]分析了动态平台迁移策略,将其建模为"领导—追随者"不完全信息博弈模型,将威胁模型构建为攻击者连续破获 k 个存在漏洞的平台,分别分析了静态和自适应攻击条件下的最优策略. 其中,前者的最优策略为最大化平台差异性,后者的最优策略为最大化平台差异性和最小化攻击者挖掘新漏洞概率之间的均衡. 然而,其构建的威胁模型过于严苛且未考虑攻防成本. 文献[9]提出一种基于操作系统动态迁移的多阶段博弈模型,指出当防御者使用多样性策略时,攻击者应选择相似性最小的操作系统进行漏洞挖掘;当防御者使用随机性策略时,攻击者应选择相似性最大的操作系统进行漏洞挖掘,该模型是一种攻击视角下的最优策略选取. 文献[10]采用贝叶斯攻击图描述攻击者利用系统漏洞之间的关联关系和节点连接关系获取和提升权限的过程,防御者以保证攻击图关键节点和最小化防御开销为优化目标. 但是,攻击图存在状态爆炸的可能性且随机攻击策略的假设与实际攻防对抗不符. 文献[11]利用改进的 FlipIt 博弈模型对不同攻防策略进行分析,并指出防御有效性依赖于攻击检测能力. 但是,博弈过程仅考虑了探测攻击(probe)和重塑防御(reimage). 文献[12]提出一种改进 FlipIt 的博弈模型,重点分析了 time-to-success 时间的概率密度函数和不同防御策略对博弈收益的影响,但对攻击者的完全信息假设和防御者的零认知能力假设不够合理. 文献[13]将 Web 应用环境下的 MTD 防御建模为贝叶斯斯坦伯格博弈模型,分析了系统漏洞重要程度和不同类型攻击者的敏感程度对攻防收益的影响,以寻求最优策略. 但是,该模型未对攻防收益进行量化分析且仅适用于 Web 软件栈环境. 文献[14]利用依赖图 DG(dependency graphs)分析了 MTD 策略对普通用户、防御者和攻击者收益的影响,将最优 MTD 策略选取抽象为寻找能够最小化用户和防御者开销且最大化攻击者开销的依赖图路径. 但是,攻防双方完全信息的假设在实际攻防对抗中难以实现,依赖图的规模存在节点状态爆炸的弊端,也未考虑攻防双方的收益. 文献[15]提出了基于马尔科夫决策的 MTD 博弈模型,对单目标和多目标 IP 跳变策略进行分析,说明了多元素跳变可以有效提高防御收益. 然而,模型构建仅针对 IP 跳变特定场景,通用性差,同时未给出具体的策略选取算法,指导性不强.

通过上述分析发现,现有最优 MTD 防御策略选取研究存在以下问题:(1)现有博弈模型研究中,攻防双方的完全信息假设以及单阶段博弈分析与实际攻防对抗不相符;(2)对 MTD 策略收益量化分析不够全面,特别是缺乏 MTD 策略动态性和多样性特征对攻防策略收益影响的分析;(3)面向 MTD 的攻防博弈模型往往针对特定的攻防策略或者面向特定的攻防环境,导致攻防策略分析不够全面,博弈模型的通用性和扩展性较差.

2　面向动态目标防御的博弈模型

MTD 环境下的攻防对抗是一种非合作、不完全信息、多阶段动态博弈过程,本文采用信号博弈对其进行分析. 其中,攻击者是信号发送方,防御者是信号接收方,攻击策略可以看作是攻击者发送的信号,防御者通过对攻击策略的分析可以推断攻击者类型并修正先验推断值,进而选取最优防御策略.

2.1　动态目标防御单阶段博弈模型

(1)博弈定义

定义 1　动态目标防御单阶段博弈模型 MS^2GM(Moving target defense Single-Stage Game Model)用七元组$(N,\Theta,A,D,P,\tilde{P},U)$描述,其中:

①$N=\{N_a,N_d\}$为博弈局中人空间. N_a 为攻击者,N_d 为防御者. 此外,还有虚拟参与人"自然".

②$\Theta=\{\Theta_a,\Theta_d\}$为博弈局中人 N_a 和 N_d 的类型空间. 其中,$\Theta_a=\{\theta_1,\theta_2,\cdots,\theta_n\}$,$n\geqslant 1$,$n\in N^+$,表示攻击者类型集合,$\Theta_a$ 的类型对于攻击者是已知的,对于防御者是随机变量,但是 Θ_a 的概率分布是共同知识;$\Theta_d=\{\theta_d\}$表示防御者只有一种类型.

③$A=\{a_1,\cdots,a_g\}$为攻击策略集,$g\geqslant 1$,$g\in N^+$.

④$D=\{d_1,\cdots,d_h\}$为防御策略集,$h\geqslant 1$,$h\in N^+$.

⑤P 为防御者先验信念集合,表示防御者对攻击者

类型的初始判断. 其中,$P=\{p_1,p_2,\cdots,p_n\}$,$p_i=p(\theta_i)>0$,$\sum_{i=1}^{n}p_i=1$,$1\leqslant i\leqslant n$.

⑥$\tilde{P}$为防御者后验信念集合,表示防御者观察到攻击策略后,使用贝叶斯法则调整对攻击者类型的推断. 其中,$\tilde{p}_{ij}=\tilde{p}(\theta_i|a_j)$,$1\leqslant i\leqslant n$,$1\leqslant j\leqslant g$.

⑦$U=\{U_a,U_d\}$为攻防御效用函数集合,分别为 $U_a=f_a(\theta_i,a_j,d_k)$和 $U_d=f_d(\theta_i,a_j,d_k)$,$1\leqslant i\leqslant n$,$1\leqslant j\leqslant g$,$1\leqslant k\leqslant h$,其描述了攻防双方从博弈当中获得的收益,不同的攻防策略具有不同的收益值.

(2)攻防收益量化

攻防策略收益量化是最优防御策略选取的基础,结合文献[16],本文从攻击面转移角度出发,对攻防策略收益进行量化分析.

定义 2 防御成本 DC(Defense Cost)包括 MTD 防御的攻击面转移成本 ASSC(Attack Surface Shifting Cost)、负面影响成本 NC(Negative Cost)和攻击识别成本 AIC(Attack Identification Cost). ASSC 指攻击面转移时改变系统资源的开销,该成本大小与预改变的系统攻击面维度相关;NC 指攻击面转移时改变系统资源,导致系统无法正常工作或服务质量下降带来的损失,该成本大小与防御策略的转移周期相关,周期越短,NC 越大;AIC 是指检测和识别不同类型攻击者的成本,能力水平越高的攻击者越难被检测和识别,其成本也越高.

$$DC(d_k,\theta_i,t^h)=ASSC(d_k)+NC(d_k,t^h)+AIC(d_k,\theta_i)$$

定义 3 攻击成本 AC(Attack Cost)指攻击者为发现和利用攻击面系统资源而付出的成本,通常包括发现和入侵系统资源时需要的时间、软硬件资源、专业知识及风险成本等. 不同能力水平的攻击者使用同一系统资源付出的成本是不一样的,能力水平越高,成本越低,表示为 $AC(a_j,\theta_i)$.

定义 4 系统损失 SD(System Damage)指攻击者使用攻击面系统资源给系统带来的损失,通常用目标资源重要程度 C(Criticality)、攻击致命度 AL(Attack Lethality)和安全属性损害 SAD(Security Attribute Damage)进行描述,详见文献[5],表示为 $SD(a_j)$. 此外,本文认为系统损失还与攻击者对攻击面系统资源的控制时间有关,即攻击者对攻击面的控制时间越长,系统损失越大.

定义 5 防御有效性 DE(Defense Effectiveness)指防御者实施防御策略引起攻击面转移后对攻击者收益的影响. 如果攻击者已经获取对系统资源的控制权限,攻击面转移消除了攻击者的既有权限;如果攻击者还未获取到对系统资源的控制权限,攻击面转移增加了攻击者获取控制权限的难度.

防御有效性有 2 个特点:(1)同一防御策略对不同攻击策略的有效性不同;(2)同一防御策略在不同类型攻击者实施同一攻击策略下的有效性不同. 不同攻防策略对的有效性可以用防御有效性函数矩阵 $\boldsymbol{E}(\theta_i)$表示,记为

$$\boldsymbol{E}(\theta_i)=\begin{bmatrix} & d_1 & d_2 & \cdots & d_h \\ a_1 & g(t,\pi_i^{11}) & g(t,\pi_i^{12}) & \cdots & g(t,\pi_i^{1h}) \\ a_2 & g(t,\pi_i^{21}) & g(t,\pi_i^{22}) & \cdots & g(t,\pi_i^{2h}) \\ \vdots & \vdots & \vdots & g(t,\pi_i^{jk}) & \vdots \\ a_g & g(t,\pi_i^{g1}) & g(t,\pi_i^{g2}) & \cdots & g(t,\pi_i^{gh}) \end{bmatrix}$$

其中,$g(t,\pi_i^{jk})$表示防御策略 $d_k(1\leqslant k\leqslant h)$对类型 $\theta_i(\theta_i\in\Theta_a)$攻击者的攻击策略 $a_j(1\leqslant j\leqslant g)$的防御有效性函数,$0\leqslant g(t,\pi_i^{jk})\leqslant 1$,$\partial g(t,\pi_i^{jk})/\partial t>0$,$t>0$,选取 sigmoid 函数度量有效性,设 $g(t,\pi_i^{jk})$为

$$g(t,\pi_i^{jk})=\frac{1-e^{-\pi_i^{jk}t}}{1+e^{-\pi_i^{jk}t}}$$

其中,π_i^{jk}为 $g(t,\pi_i^{jk})$的最大一阶导数,攻击者能力越强,π_i^{jk}越大. π_i^{jk}是防御有效性函数矩阵的关键因素,具有以下特点:(1)不失一般性,设 $\theta_{i+1}\rhd\theta_i(\theta_i\in\Theta_a,\theta_{i+1}\in\Theta_a)$表示 θ_{i+1}类型攻击者的攻击能力强于 θ_i 类型攻击者,如果 $\theta_{i+1}\rhd\theta_i$,则在同一攻防策略 d_k 和 a_j 下,$\pi_{i+1}^{jk}>\pi_i^{jk}$,这意味着攻击能力越强的攻击者能够花费更短的时间夺取对攻击面资源的控制;(2)对于系统执行栈中处于同一层次的一类防御策略 $d_k(1\leqslant k\leqslant h)$而言,设 $\gamma(d_k,d_{k+1})$表示防御策略 d_k 和 d_{k+1}的差异性,如果 $\gamma(d_k,d_{k+1})<\gamma(d_k,d_{k+2})$,则在应对同一攻击策略 a_j 时,$\pi_i^{j(k+1)}<\pi_i^{j(k+2)}$,这意味着防御策略的差异性越小,攻击者能够花费更短的时间夺取对攻击面资源的控制. 由于系统执行栈不同层次的防御策略在组成元素和生效机理上的不同,差异性的度量指标也各不相同,需要根据实际的应用环境进行设置,例如对于运行平台层的操作系统多样化策略,可以根据系统代码相似性进行度量[9];对于网络层的 IP 跳变策略,可以根据跳变空间大小进行度量.

设 $t^*(\theta_i,a_j,d_k)\in\underset{t}{\operatorname{argmin}}(1-g(t,\pi_i^{jk}))\leqslant\delta$,即一旦攻击有效性大于预定阈值,攻击者就夺取了对系统资源的控制,则 $t^*(\theta_i,a_j,d_k)$表示攻击者在攻防策略对(a_j,d_k)下对攻击面系统资源的控制时间.

定义 6 攻击收益 AR(Attack Reward)指攻击者使用攻击面系统资源获得的收益.

$$AR(\theta_i,a_j,d_k)=\varphi(t^h,t^*)(1+g_a(t^h,t^*))SD(a_j)+DC(d_k,\theta_i,t^h)-AC(a_j,\theta_i)$$

$$\varphi(x,y)=\begin{cases}1, & x\geqslant y\\ 0, & x<y\end{cases},x\in\boldsymbol{R},y\in\boldsymbol{R}$$

$g_a(t^h,t^*)$表示攻击者对攻击面系统资源的控制时间对攻击收益的影响,是 t^h-t^* 的单调增函数. 其中,t^h

$-t^*$ 表示攻击者控制攻击面系统资源的时间.

定义 7　防御收益 DR(Defense Reward)指防御者转移攻击面系统资源获得的收益.

$$DR(\theta_i,a_j,d_k)=g_d(t^h,t^*)SD(a_j)+AC(a_j,\theta_i)-DC(d_k,\theta_i,t^h)$$

$g_d(t^h,t^*)$ 表示防御者对攻击面系统资源控制时间对防御收益的影响,是 $\min(t^h,t^*)$ 的单调增函数.

(3)博弈顺序

虚拟参与人"自然"赋予攻击者类型一个先验概率,该概率值可由历史经验获得,也可以平均分配概率. 防御者通过观察攻击行为不断修正对攻击者类型的信念推断,调整防御策略. 博弈顺序如下:

①"自然"按一定概率从攻击者 N_a 的类型空间 Θ_a 中选择一个类型 θ_i,其中 $\theta_i\in\Theta_a$. 攻击者 N_a 知道 θ_i,防御者 N_d 不知道,但防御者 N_d 拥有对 θ_i 的推断,即知道攻击者类型的先验概率.

②攻击者 N_a 在观察到 θ_i 后从其策略空间 A 中选择一条攻击策略 a_j.

③防御者 N_d 在观察到 a_j 后,先应用贝叶斯法则从先验概率得到后验概率,再从其策略空间中 D 选择一条防御策略 d_k.

④攻防双方的收益分别由 $U_a=f_a(\theta_i,a_j,d_k)$ 和 $U_d=f_d(\theta_i,a_j,d_k)$ 计算得出.

在2种情况下博弈结束:1)防御者所采取的防御策略能够完全堵截各条攻击路径,2)攻击者已经成功到达攻击目标.

(4)S^2MGM 精炼贝叶斯均衡求解

①防御者在每个信息集上建立后验概率推断 $\tilde{p}(\theta_i|a_j)$.

②防御者推断最优防御策略 $d^*(a_j)$.

③攻击者推断最优攻击策略 $a^*(\theta_i)$.

④精炼贝叶斯均衡解$(a^*(\theta),d^*(a),\tilde{p}^*(\theta_i|a_j))$.

利用②、③两步所得的子博弈精炼纳什均衡$(a^*(\theta),d^*(a))$,求出满足贝叶斯法则的防御者对攻击者类型的推断 $\tilde{p}_{ij}^*=\tilde{p}^*(\theta_i|a_j)$,如果 $\tilde{p}(\theta_i|a_j)$ 与 $\tilde{p}^*(\theta_i|a_j)$ 不冲突,则$(a^*(\theta),d^*(a),\tilde{p}^*(\theta_i|a_j))$为精炼贝叶斯均衡.

MS^2GM 博弈考虑了攻防双方的信息不完全性和攻防行动的序贯特性,以及攻防行为对对方策略选取的影响,与完全信息和静态信息博弈相比,更加符合实际攻防场景,具有更好的准确性和适用性.

2.2　动态目标防御多阶段博弈模型

随着攻击者和防御者策略选择的交互进行,在每一个连续的时间周期 $T_t,t=1,2,\cdots,n(n\in Z^+)$,$MS^2GM$ 博弈将被重复进行. 为简化起见,假设攻防双方在 T_t 和 T_{t-1} 阶段博弈具有相同的收益,即多阶段动态目标防御博弈中不存在收益的折扣现象.

根据贝叶斯法则,防御者可以通过 T_{t-1} 阶段博弈更新 T_t 阶段博弈对攻击者类型的推断. 设 $h_a(T_t)$ 为攻击者的在 T_t 阶段之前的历史策略集合,$a_j(T_t)$ 为攻击者在 T_t 阶段博弈的动作.

$\tilde{p}(\theta_i|a_j(T_t),$

$$h_a(T_t))=\frac{p(\theta_i\mid h_a(T_t))p(a_j(T_t)\mid\theta_i,h_a(T_t))}{\sum_{\theta_i\in\Theta_a}p(\theta_i\mid h_a(T_t))p(a_j(T_t)\mid\theta_i,h_a(T_t))}$$

其中,$p(\theta_i|h_a(T_t))$ 表示在历史策略集合 $h_a(T_t)$ 下的先验推断,$p(a_j(T_t)|\theta_i,h_a(T_t))$ 表示在 T_t 阶段博弈中攻击者在采取历史策略集合 $h_a(T_t)$ 的前提下选取策略 $a_j(T_t)$ 的概率.

定义 8　动态目标防御多阶段博弈模型 M^2SGM(Moving target defense Multi-Stage Game Model)用七元组$(N,\Theta,A,D,P,\tilde{P},U)$描述,其中:

①N,Θ,A,D,P,U 的定义与 MS^2GM 模型中的定义相同.

②$\tilde{p}_{ij}=\tilde{p}(\theta_i|a_j(T_t),h_a(T_t))$ 表示攻击者在 T_t 阶段采取历史策略集合 $h_a(T_t)$ 的前提下,类型为 θ_i 的概率,在 T_t 阶段博弈结束时,对推断值 $\tilde{p}_{ij}=\tilde{p}(\theta_i|a_j(T_t),h_a(T_t))$ 进行更新.

随着推断值的更新,M^2SGM 博弈以序贯方式进行,最后通过贝叶斯均衡表示 M^2SGM 均衡. 在整个博弈过程中,攻防双方为最大化收益,并不总是在每一个阶段博弈中采取相同策略,而且随着博弈的进行,双方的最优响应策略与可能改变的推断值独立. 在寻找 M^2SGM 博弈的贝叶斯均衡解之前,首先说明该博弈模型满足必须的贝叶斯条件.

定义 9　贝叶斯条件. 贝叶斯条件包括:

C_1:"后验推断"是相互独立的,并且参与者的所有类型具有相同的"先验推断";

C_2:"先验推断"到"后验推断"的更新通过贝叶斯规则实现;

C_3:参与者不传递任何参与者所不知道的信号;

C_4:"后验概率"在 Θ 上的共同的联合概率分布是一致的.

引理 1　M^2SGM 博弈满足贝叶斯条件.

证明　(1)防御者有且仅有一种类型,C_1 满足. (2)式 $\tilde{p}(\theta_i|a_j(T_t),h_a(T_t))$ 由贝叶斯规则得到,C_2 满足. (3)攻击者的信号由其自身策略决定,并且如果 $a_j(T_t)=a'_j(T_t)$ 成立,那么 $p(a_j(T_t)|\theta_i,h_a(T_t))=p(a'_j(T_t)|\theta_i,h_a(T_t))$,$C_3$ 满足. (4)现实网络攻防对抗过程中,通常是多个攻击者对一个防御目标发起协同攻击,但是为了简化分析,将博弈过程中的多个攻击者抽象为一个,即 M^2SGM 博弈在任何阶段都只有两个参与者,C_4 满足.

对攻击者类型的推断值进行更新后,M^2SGM 博弈

的精炼贝叶斯均衡求解方法与 MS^{2}GM 博弈类似.

2.3 博弈模型特点总结

(1)攻防策略收益量化更加合理. 本文以攻防双方抢占和控制构成攻击面的系统资源的成本和收益为出发点,通过引入防御策略转移周期、防御策略多样性对防御有效性的影响等因素,以 MTD 策略自身特点为基础,分析了攻防收益量化,可保证后续博弈均衡解的有效性和准确性.

(2)博弈模型更加符合实际攻防环境. 与完全信息博弈相比,增加了对攻击者类型不确定性的考虑;与静态博弈相比,增加了对攻防行为非同时性的考虑;与单阶段博弈相比,增加了对攻防对抗多回合性的考虑. 特别是,防御者采用后验概率修正对攻击者类型的推断,符合每一回合的攻防博弈结束后防御者利用数据统计和异常检测等方式获取攻击者信息,进而动态修正对攻击者类型判断的特点.

(3)博弈模型的通用性和扩展性更好. 现有博弈模型往往针对单一攻防类型或特定攻防策略,如文献[14]只针对动态平台策略、文献[16]只针对攻击检测策略,其适用范围有限且难以推广扩展,M^{2}SGM 模型适用于具有多种攻击类型和攻防策略的 MTD 环境,通用性和扩展性更好.

3 仿真实验与分析

3.1 仿真实验环境描述

为验证上述博弈模型以及均衡求解方法的正确性和有效性,借鉴 Sheyner 构建的典型实验网络[17],设计了实验目标网络,拓扑结构如图 1 所示.

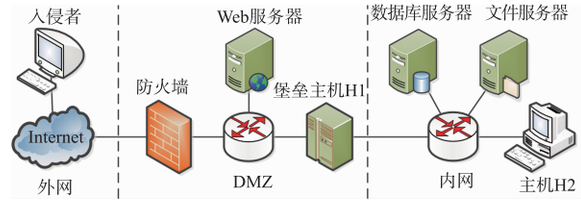

图1 实验目标网络拓扑结构

安全威胁来自外网,防火墙的安全策略为仅允许外部用户访问 Web 服务器的 HTTP 和 FTP 服务,以及堡垒主机 H_1 上的 SMTP 服务;访问控制规则规定 DMZ 的 Web 服务器和内网的主机 H_2 可以对文件服务器和数据库服务器进行访问;其他网络节点和端口均进行阻断. 攻击者初始时刻无法直接访问数据库服务器,其目的是通过一系列的原子攻击获取对数据库服务器的 root 访问权限. 使用 Nessus 工具进行扫描,挖掘设备脆弱性信息,如表 1 所示.

表 1 设备脆弱性信息

序号	主机	服务	脆弱性	CVE 编号	初始权限	目标权限
1	Web 服务器	IIS	HTTP. sys	CVE-2015-1635	user	root
2	Web 服务器	FTP	Write $ home/. rhost	CVE-2011-4800	user	user
3	堡垒主机 H_1	SMTP	Remote buffer overflow	CVE-2005-0560	user	root
4	主机 H_2	RPC	Code Injection	CVE-2008-4250	user	root
5	文件服务器	FTP	Write $ home/. rhost	CVE-2011-4800	user	user
6	数据库服务器	FTP	Write $ home/. rhost	CVE-2011-4800	user	user
7	数据库服务器	Mysql	Local buffer overflow	CVE-2005-2558	user	root

攻击者首先通过扫描攻击探测到 Web 服务器的 IP 地址及其提供的 IIS 服务和 FTP 服务存在的漏洞,利用 CVE-2015-1635 漏洞获取 Web 服务器的 root 权限,或者利用 CVE-2011-4800 漏洞获取 Web 服务器的 user 权限;然后,利用堡垒主机 H_1 的 CVE-2005-0560 漏洞获取对其的 root 权限;再次,利用内网主机 H_2 的 CVE-2008-4250 漏洞获取对其的 root 权限,或者利用文件服务器的 CVE-2011-4800 漏洞获取对其的 user 权限;最后,利用数据库服务器的 CVE-2005-2558 漏洞获取对其的 root 权限. 为简洁有效地说明问题,仅对攻击者试图获取 Web 服务器控制权限这一阶段进行博弈分析,实例中将攻击者分为高水平和低水平攻击者 2 类,攻击者可进行的原子攻击策略描述如表 2 所示.

表 2 原子攻击策略描述

攻击策略	策略描述	攻击类别	攻击致命度	攻击成本	
				低水平	高水平
a_1	Network scanning	probe	2	10	8
a_2	Network sniffer	probe	3	15	12
a_3	Code injection	root	10	25	20
a_4	Remote buffer overflow	root	10	35	30
a_5	ftp. rhost	user	5	15	12
a_6	Local buffer overflow	root	10	20	15

当前 Web 服务器部署的 Web 服务软件栈为“Windows 7 SP1 + IIS 7.0”. Web 服务器上可实施的 MTD 策略如表 3 所示,采取固定跳变周期方式,地址跳变策略

指改变 Web 服务器 IP 地址;Web 服务架构多态化策略[7]是指通过软件栈实现 Web 服务,软件栈通常包括 Web 服务器程序、Web 应用程序、操作系统和虚拟层,其具体实现是通过建立一系列多态化虚拟服务器,每个服务器配置一个唯一的软件栈,以建立多态化的攻击面,例如"IIS 服务器 + Windows"和"Apache 服务器 + Ubuntu"是在 Web 服务器程序和操作系统层面的两种不同实现方式.

表 3　动态目标防御策略描述

防御策略	策略描述	跳变周期	操作成本	负面成本
d_1	Web 服务器地址跳变	30	10	15
d_2	Web 服务器地址跳变	50	10	10
d_3	Windows 8 + IIS 7.0	100	20	20
d_4	Ubuntu 14.04 + Apache2	100	20	20

3.2　收益计算

图 1 所示的实验网络拓扑中,第一阶段的攻防博弈将在攻击者试图获取对 Web 服务器的控制权限时展开. 在此,对攻击策略为 a_1 和 a_2、防御策略为 d_1、d_2、d_3、d_4 情形下的攻防收益进行量化分析.

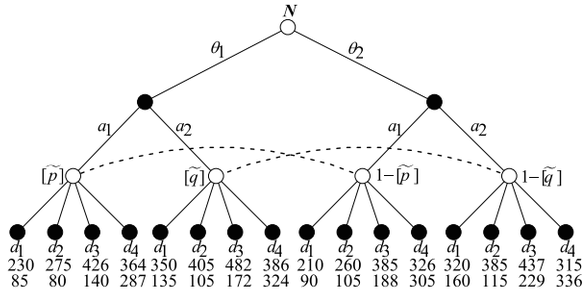

图2　动态目标防御环境下的第一阶段博弈树

设防御者对高水平和低水平类型攻击者的先验信念为$(\theta_1,\theta_2)=(0.5,0.5)$,收益量化参数为 $\delta=0.02$,$\pi_1^{11}=\pi_1^{12}=0.18$,$\pi_1^{13}=0.2$,$\pi_1^{14}=0.4$,$\pi_1^{21}=\pi_1^{22}=0.2$,$\pi_1^{23}=0.3$,$\pi_1^{24}=0.5$,$\pi_2^{11}=\pi_2^{12}=0.15$,$\pi_2^{13}=0.15$,$\pi_2^{14}=0.25$,$\pi_2^{21}=\pi_2^{22}=0.15$,$\pi_2^{23}=0.2$,$\pi_2^{24}=0.35$,$g_a(t^h,t^*)=(t^h-t^*)/t^h$,$g_d(t^h,t^*)=\min(t^*,t^h)/t^h$. Web 服务器、文件服务器和数据库服务器的安全属性代价为 20、20 和 30,重要性程度为 4、4 和 5. 防御者对高、低水平攻击者的信号分析成本分别为 15 和 10. 利用攻防收益量化公式计算攻防收益,得到动态目标防御环境下的第一阶段博弈树,如图 2 所示.

3.3　均衡求解和防御策略选取

网络攻防对抗中,防御者通常会以多种防御策略的不同组合方式进行防御,以达到最大化防御收益的效果,依循现有不完全信息动态博弈求解精炼贝叶斯均衡解的过程即可求解到最优防御策略组合. 但是,多种 MTD 策略的混合均衡解难以体现出不同跳变对象、不同跳变周期 MTD 策略对博弈均衡解的影响. 为此,本节将图 2 所示的博弈树进行拆分,分别考虑表 3 中不同 MTD 策略要素对博弈均衡解的影响. 具体而言,一是考虑攻击策略为 a_1 和 a_1、防御策略为 d_1 和 d_2 的博弈,以分析不同跳变周期、无差异性跳变对象对博弈均衡解的影响;二是考虑攻击策略为 a_1 和 a_1、防御策略为 d_3 和 d_4 的博弈,以分析相同跳变周期、差异性跳变对象对博弈均衡解的影响. 对于前者,本节给出具体的求解过程;对于后者,在 3.5 节给出其一般性分析结论.

假设攻击者有 4 种纯策略:"自然"选择类型 θ_1,攻击者选择动作 a_1,自然选择类型 θ_2,攻击者选择动作 a_1,该策略记为(a_1,a_1);同理可得(a_1,a_2)、(a_2,a_1)和(a_2,a_2). 防御者也有 4 种纯策略:(d_1,d_1),(d_1,d_2),(d_2,d_1),(d_2,d_2). 在此,以混同均衡策略(a_2,a_2)和分离均衡策略(a_2,a_1)为例进行求解.

混同均衡策略(a_2,a_2)

对于攻击策略(a_2,a_2),攻击者的均衡策略为 $a^*(\theta_1)=a^*(\theta_2)=a_2$,由图 2 知防御者右边用虚线连接的信息集处于均衡路径之上,得出后验概率 $\tilde{q}(\theta_1\mid a_2)=\tilde{q}(\theta_2\mid a_2)=0.5$,则防御者采取策略 d_1 时的收益为 147.5,防御者采取策略 d_2 时的收益为 110. 因此,防御者的最优策略是 $d^*(a_2)=d_1$. 此时,类型为 θ_1 和 θ_2 的攻击者可得到的收益分别为 350 和 320.

为确认 $a^*(\theta_1)=a^*(\theta_2)=a_2$ 是攻击者的最优选择,还需要检查攻击者采取策略 a_1 时的情形. 若攻击者采取策略 a_1,防御者观测到 a_1,博弈进入图 2 左边用虚线连接的信息集.

如果防御者对 a_1 的反应为 d_1,则类型为 θ_1 和 θ_2 的攻击者选择 a_1 的收益分别为 230 和 210,选择 a_2 的收益分别为 350 和 320,此时,类型为 θ_1 和 θ_1 的攻击者都会选择策略 a_2;如果防御者对 a_1 的反应为 d_2,则类型为 θ_1 和 θ_2 的攻击者选择 a_1 的收益分别为 275 和 260,而选择 a_2 的收益分别为 350 和 320,此时,类型为 θ_1 和 θ_2 的攻击者仍会选择 a_2.

此外,还需考虑防御者对应于 a_1 的信息集中的推断$[\tilde{p},1-\tilde{p}]$,以及给定这一推断时选择策略 d_1 是否是最优的. 防御者采取策略 d_1 的收益为

$$\sum_{\theta_i\in\Theta_a}\tilde{p}(\theta_i\mid a_1)f_d(\theta_i,a_1,d_1)=\tilde{p}\times 85+(1-\tilde{p})\times 90=90-5\tilde{p}$$

防御者采取策略 d_2 时的收益为

$$\sum_{\theta_i\in\Theta_a}\tilde{p}(\theta_i\mid a_1)f_d(\theta_i,a_1,d_2)=\tilde{p}\times 80+(1-\tilde{p})\times 105=105-25\tilde{p}$$

当 $0.75\leqslant\tilde{p}\leqslant 1$ 时,$90-5\tilde{p}>105-25\tilde{p}$,防御者采取策略 d_1 是最优的.

所以,$\{(a_2,a_2),(d_1,d_1);0.75\leqslant\tilde{p}\leqslant 1;\tilde{q}=0.5\}$ 为该博弈的混同精炼贝叶斯均衡.

分离均衡策略(a_2,a_1)

防御者对应于a_1的信息集处在均衡路径上的贝叶斯推断为(0,1),防御者对应于攻击策略a_2的信息集处在均衡路径上的贝叶斯推断为(1,0). 防御者在此推断下的最优反应分别为d_2和d_1,攻击者的收益分别为260和350. 此时,需要确定给定的防御者策略(d_1,d_2),攻击者不会偏离策略(a_2,a_1).

如果类型为θ_1的攻击者选择a_1,防御者的反应为d_1,攻击者的收益(210<260)将减小;如果类型为θ_2的攻击者选择a_2,防御者的反应为d_1,攻击者的收益(320<350)也将减小. 因此,两种类型的攻击者没有任何动机偏离策略(a_2,a_1).

所以,$\{(a_2,a_1),(d_2,d_1);\tilde{p}=0;\tilde{q}=1\}$为该博弈的分离精炼贝叶斯均衡.

同理,求得混同均衡策略(a_1,a_1)和分离均衡策略(a_1,a_2)均不能构成该博弈的精炼贝叶斯均衡.

3.4 修正先验推断

通过对T_t时间之前的历史统计得到不同类型攻击者使用不同攻击策略的概率$p(a_j|\theta_i,h_a(T_t))$,如表4所示.

表4 攻击类型与攻击策略概率关系

$p(a_j\|\theta_i,h_a(T_t))$	a_1	a_2	a_3	a_4	a_5	a_6
θ_1	0.695	0.758	0.526	0.694	0.517	0.795
θ_2	0.762	0.728	0.547	0.525	0.682	0.496

当在第T_t时间段开始时,攻击者采取攻击策略a_1,可得到攻击者类型的后验概率为$\tilde{p}(\theta_1|h_a(T_t))=0.52$,$\tilde{p}(\theta_2|h_a(T_t))=0.48$. 如果防御者发现攻击者在下一阶段博弈中选取攻击策略a_5,立即修正对攻击类型的推断.

$$\tilde{p}(\theta_1|h_a(T_{t+1}))=0.451<\tilde{p}(\theta_1|h_a(T_t))$$
$$\tilde{p}(\theta_2|h_a(T_{t+1}))=0.549>\tilde{p}(\theta_2|h_a(T_t))$$

由此可见,如果攻击者持续采用低水平攻击类型的攻击策略,防御者可以通过不断修正信念,增加对其为低水平攻击者的推断. 同时,对攻击者类型推断值的修正将影响下一阶段博弈的攻防收益值,进而影响到攻防双方的策略选择,这也体现了多阶段动态博弈中防御者通过不断观察攻击者行为,增加对攻击者类型的掌握,进而改变己方策略选择的优势.

3.5 实验结果分析

通过对面向MTD的攻防博弈模型均衡解和攻防收益的分析,可以得出利用MTD策略进行网络防御的一般性规律:

(1)应对攻击水平越高的攻击者,MTD策略的跳变周期应该减小. 因为攻击水平越高的攻击者发现和入侵攻击面系统资源所需的时间越短,对攻击面系统资源的控制时间也越长,带来的损害也会大于防御策略跳变引起的负面影响成本,而减小防御策略跳变周期能够有效减小攻击者对攻击面系统资源的控制时间,在很大程度上减小攻击者收益.

(2)MTD策略的差异性越大,其防御有效性越好,防御收益越高. 图2的攻防策略收益可以看出,无论是高水平攻击者还是低水平攻击者,防御策略d_4的防御收益都要明显大于防御策略d_3的防御收益. 这是因为防御策略的差异性越大,系统软件栈配置信息存在潜在威胁漏洞的相似性越小,攻击者需要更长的时间进行扫描探测和漏洞挖掘,从而提高了防御收益.

(3)防御者对攻击者类型推断值的准确性对攻防信号博弈走势和均衡解具有重大影响. 防御者不能完全掌握和判断攻击者类型,但能根据检测到的攻击行为,使用贝叶斯法则修正对攻击者类型的判断,并选择最优防御策略. 其中,防御者对攻击者类型的推断值直接影响攻防收益值的计算,进而影响攻防博弈的均衡求解. 因此,MTD策略应该与攻击检测策略共同部署,以达到最优的防御效果.

4 结束语

本文基于不完全信息动态博弈研究了最优MTD策略选取问题. 通过利用信号博弈理论,构建了面向MTD的单阶段和多阶段博弈模型,分析了博弈模型特点,给出了MTD防御的一般性规律,对于最优MTD策略选取具有一定的指导意义. 但是,MTD防御环境下的攻防策略收益量化,特别是MTD策略有效性函数矩阵,还需进行大量实验才能使量化值更加合理准确. 同时,需将MTD策略与其他防御策略相结合,以最大化综合防御收益.

参考文献

[1] ZHUANG Rui, ZHANG Su, DELOACH S A, OU Xinming, SINGHAL A. Simulation-based approaches to studying effectiveness of moving-target network defense[A]. Proceedings of the National Symposium on Moving Target Research[C]. Annapolis: DataStream Content Solutions, 2012. 15111-15126.

[2] NITRD. Cybersecurity game-change research & development recommendations [OL]. http://www.nitrd.gov/pubs/CSIA_IWG_%20Cybersecurity_%20GameChange_R D_%20Recommendations_20100513.pdf,2014-03-21.

[3] XU Jun, GUO Pin-yao, ZHAO Ming-yi, Erbacher R F, ZHU Ming-hui, LIU Peng. Comparing different moving target defense techniques[A]. Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security[C]. Arizona: ACM, 2014. 97-107.

[4] 姜伟,方滨兴,田志宏,张宏莉. 基于攻防博弈模型的网络安全测评和最优主动防御[J]. 计算机学报,2009,32(4):817 - 827.
JIANG Wei, FANG Bin-xing, TIAN Zhi-hong, ZHANG Hong-li. Evaluating network security and optimal active defense based on attack-defense game model[J]. Chinese Journal of Computers, 2009, 32(4): 817 - 827. (in Chinese)

[5] COLBAUGH R, GLASS K. Predictability-oriented defense against adaptive adversaries[A]. Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)[C]. Seoul: IEEE, 2012. 2721 - 2727.

[6] MANADHATA P K. Game Theoretic Approaches to Attack Surface Shifting[M]. New York: Moving Target Defense II: Application of Game Theory and Adversarial Modeling, Springer, 2013. 1 - 13.

[7] ZHU Quan-yan, BASAR T. Game-Theoretic Approach to Feedback-Driven Multi-Stage Moving Target Defense [A]. International Conference on Decision and Game Theory for Security[C]. Springer, 2013. 246 - 263.

[8] CARTER K M, RIORDAN J F, OKHRAVI H. A game theoretic approach to strategy determination for dynamic platform defenses[A]. Proceedings of the First ACM Workshop on Moving Target Defense [C]. Arizona: ACM, 2014. 21 - 30.

[9] WINTERROSE M L, CARTER K M, WAGNER N, STREILEIN W W. Adaptive attacker strategy development against moving target cyber defenses[J]. Computer Science, 2014, 1472: 1 - 11.

[10] MIEHLING E, RASOULI M, TENEKETZIS D. Optimal defense policies for partially observable spreading processes on Bayesian attack graphs[A]. Proceedings of the Second ACM Workshop on Moving Target Defense[C]. Denver: ACM, 2015. 67 - 76.

[11] PRAKASH A, WELLMAN M P. Empirical game-theoretic analysis for moving target defense[A]. Proceedings of the Second ACM Workshop on Moving Target Defense [C]. Denver: ACM, 2015. 57 - 65.

[12] JONES S, OUTKIN A, GEARHART J, et al. Evaluating moving target defense with PLADD[R]. Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2015.

[13] VADLAMUDI S G, SENGUPTA S, TAGUINOD M, et al. Moving target defense for web applications using Bayesian Stackelberg games[A]. Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems[C]. Singapore: International Foundation for Autonomous Agents and Multiagent Systems, 2016. 1377 - 1378.

[14] HAMLET J R, LAMB C C. Dependency graph analysis and moving target defense selection[A]. Proceedings of the 2016 ACM Workshop on Moving Target Defense [C]. Vienna: ACM, 2016. 105 - 116.

[15] MALEKI H, VALIZADEH S, KOCH W, BESTAVROS A, DIJK M V. Markov modeling of moving target defense games[A]. Proceedings of the 2016 ACM Workshop on Moving Target Defense [C]. Vienna: ACM, 2016. 81 - 92.

[16] VAN LEEUWEN B, STOUT W, URIAS V. Operational cost of deploying moving target defenses defensive work factors[A]. Proceedings of the 2015 IEEE Military Communications Conference [C]. Florida: IEEE, 2015. 966 - 971.

[17] SHEYNER O, HAINES J, JHA S, LIPPMANN R, WING J M. Automated generation and analysis of attack graphs [A]. Proceedings of 2002 IEEE Symposium on Security and Privacy[C]. California: IEEE, 2002. 273 - 284.

作者简介

刘　江　男. 1988 年 6 月出生,陕西城固人,现为信息工程大学博士研究生,主要研究方向为移动目标防御、软件定义网络安全.
E-mail: liujiang2333@163.com

张红旗　男. 1962 年 11 月出生,河北唐山人,信息工程大学教授、博士生导师,主要研究方向为信息系统安全、网络安全管理.
E-mail: zhq37922@126.com

刘　艺　女. 1991 年 12 月出生,江西崇义人,现为信息工程大学博士研究生,主要研究方向为软件定义网络安全、网络虚拟化.