

基于不完全信息随机博弈的防御决策方法

杨峻楠^{1,2}, 张红旗^{1,2}, 张传富^{1,2}

(1. 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: 现有防御决策中的随机博弈模型大多由矩阵博弈与马尔可夫决策组成, 矩阵博弈中假定防御者已知攻击者收益, 与实际不符。将攻击者收益的不确定性转换成对攻击者类型的不确定性, 构建了由静态贝叶斯博弈与马尔可夫决策结合的不完全信息随机博弈模型, 给出了不完全信息随机博弈模型的均衡求解方法, 使用稳定贝叶斯纳什均衡指导防御者的策略选取。最后通过一个具体实例验证了模型的可行性和有效性。

关键词: 网络攻防; 随机博弈; 不完全信息; 贝叶斯纳什均衡; 防御决策

中图分类号: TP393.08

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018065

Defense decision-making method based on incomplete information stochastic game

YANG Junnan^{1,2}, ZHANG Hongqi^{1,2}, ZHANG Chuanfu^{1,2}

1. Information Engineering University, Zhengzhou 450004, China
2. Henan Province Key Laboratory of Information Security, Zhengzhou, 450001, China

Abstract: Most of the stochastic game models to select network defense strategies are composed of matrix game and Markov decision, which assumes that the defender has known the attacker's revenue. This assumption does not conform to the actual situation. The uncertainty of the attacker's income was converted into the indeterminacy of the attacker type, and an incomplete information stochastic game model, which was combined with the static Bias game and Markov decision, was constructed. The equilibrium solution method of the incomplete information stochastic game model was given, and the strategy selection of the defender was guided by the stable Bias Nash equilibrium. Finally, a practical example was given to demonstrate the feasibility and effectiveness of the model.

Key words: network attack-defense, stochastic game, incomplete information, bayesian nash equilibrium, defense strategies

1 引言

随着信息化程度的不断加强, 网络给人们提供了更多便利, 但是网络攻击也日趋频繁, 给受攻击者造成巨大损失^[1]。政府、企业和学校等机构每年

都要在网络安全上投入大量资金, 但是由于网络本身的复杂性以及当前技术的限制, 导致无法保证网络绝对的安全, 亟需一种技术能够对攻防行为进行分析, 对网络风险与安全投入进行有效折中, 使防御者能够利用有限的资源做出合理的决策。博弈论

收稿日期: 2018-06-03; 修回日期: 2018-07-11

通信作者: 杨峻楠, 624519905@qq.com

基金项目: 国家高技术研究发展计划基金资助项目 (“863” 计划) (No.2014AA7116082, No.2015AA7116040)

Foundation Item: The National High Technology Research and Development Program of China (863 Program) (No.2014AA7116082, No.2015AA7116040)

与网络攻防所具有的目标对立性、关系非合作性和策略依存性高度契合^[2]。基于博弈论的网络安全研究逐渐成为一个热点^[3]。

目前,国内外基于博弈论的网络攻防行为分析与防御策略选已取得较多成果。很多学者使用单阶段博弈对网络攻防进行分析。文献[4]基于静态博弈理论对蠕虫病毒攻防行为进行分析,采用贝叶斯纳什均衡对防御策略进行指导。文献[5]基于动态博弈理论研究DDoS的攻防行为并依据精练纳什均衡设计了最优防御策略选取算法。单阶段博弈能够在一定程度上指导网络防御,但其只能分析一个攻防回合,而攻防对抗往往由多阶段组成,为了能够更加符合实际,部分学者开始使用多阶段博弈对网络攻防进行建模分析。文献[6]构建多阶段攻防信号博弈模型研究最优主动防御策略选取问题,并给出了博弈均衡求解方法。文献[7]针对无线传感器网络中的安全防御问题,构建了多阶段重复博弈模型。上述成果虽然可以对攻防对抗的多个回合进行分析,但是只考虑了攻防双方对网络对抗的影响,忽略了网络本身对攻防进程的影响。网络状态的转移不仅受到攻防动作的影响,还受其他一些因素的影响,其状态转移呈现随机性。上述成果没有对这种随机性进行分析,削弱了其指导价值。随机博弈^[8]是一种多阶段博弈,其采用马尔可夫过程描述博弈系统的状态转移,能够准确刻画网络攻防对抗过程。文献[9]将攻击者的特权状态作为网络状态构建完全信息随机博弈模型,对攻防的收益量化进行分析并依据纳什均衡提出了防御策略选取算法。文献[10]使用随机博弈对网络对抗问题进行研究,运用凸分析理论证明均衡的存在性并将均衡求解转换为一个非线性规划问题。Wei等^[11]利用完全信息随机博弈分析攻击者与防御者之间的策略交互,利用纳什均衡评估不同状态下系统面临的风险。Ding等^[12]在完全信息随机博弈的框架下,研究了传感器与DoS攻击者之间的相互作用过程,根据纳什均衡给出了传感器的最优功率方案。上述研究的随机博弈由矩阵博弈与马尔可夫决策组成,使用马尔可夫决策描述网络的演进,使用矩阵博弈分析各回合攻防双方的决策。由于攻防双方的目标对立性,完全掌握对手信息几乎不可能,而上述

成果中使用的矩阵博弈以完全信息假设为前提,与实际不符。

基于上述分析,为解决网络防御决策问题,将防御者对攻击者收益的不确定转换成对攻击者类型的不确定,构建由静态贝叶斯博弈与马尔可夫决策结合的不完全信息随机博弈模型,针对博弈模型给出了一种非线性规划求解方法,依据稳定贝叶斯纳什均衡对攻击行为进行预测和对防御策略进行指导。

本文的主要贡献如下。

- 1) 从攻防策略的角度将复杂的网络攻防问题抽象为一个双人、非零和的不完全信息随机博弈问题。
- 2) 借鉴现有的完全信息随机博弈求解方法,给出了本文不完全信息随机博弈模型求解方法。
- 3) 使用本文方法对一个典型的网络攻防场景进行了分析,提供了一个详细的实例。

2 不完全信息随机博弈及防御策略选取

2.1 问题描述与分析

网络攻防对抗是一个复杂问题,本文主要从决策角度对其进行研究。网络攻防对抗包含3个主要对象,分别是攻击者、防御者和网络本身。攻防双方依据自己的策略选择攻防动作,网络在攻防双方的联合行为作用下进行演进。由于网络中一些未知因素的影响,导致攻防演进呈现不确定性。网络对抗中攻防双方都是理性的,双方的目的是获得最大收益,这是双方制定攻防策略的准则。

本文的研究对象是企业、政府和学校等局域网,网络管理员作为防御者负责全网的防御措施,攻击者来自外网,远程对局域网发起攻击。在攻防过程中,攻防角色不会互换,即防御者仅采取防御动作,不会以攻代防。

2.2 不完全信息随机博弈模型

使用随机博弈对网络攻防过程进行建模分析。网络攻防是一个连续过程,对其进行离散化处理,将整个时间轴划分为一个个时间片,每个时间片有且只有一个网络状态,攻防双方在每个网络状态依据攻防策略选择攻防动作,网络在攻防双方的联合行为作用下由一个状态转移到另一

个状态。网络状态独立于时间，即不同的时间片内，可能有相同的状态。攻防策略与网络状态相关联，而与时间片无关。使用随机博弈对网络攻防建模主要分为2个部分。

1) 建立每个网络状态的博弈模型

在实际中，同一时刻可能有不止一个攻击者攻击网络，但是对于防御者来说，这些攻击者是无差别的，防御者可以将不同攻击者当作是一个集体，将不同攻击者的行为看作是同一个攻击者的不同行为，这样，网络攻防对抗就可以使用一个双人博弈进行分析。

攻击者的收益受攻击目标、攻击者能力等多方面因素的影响，又由于攻防双方是非合作关系，导致最终防御者往往无法确定攻击收益函数，即攻击收益对防御者是未知信息，所以网络攻防是一个不完全信息博弈。本文使用海萨尼转换^[13]对其进行处理，将防御者对攻击者收益的不确定性转化为对攻击者类型的不确定性，攻击者类型为攻击者的私有信息，但是防御者对攻击者类型分布有一个概率判断，这个概率判断为共同知识。攻防双方的收益受多种因素影响，两者之间并不是简单的零和，故非零和博弈更符合攻防对抗实际。

由于攻防双方的非合作性，攻防双方只能通过检测网络了解对手行动，这往往比动作的执行时间有所延迟，所以在每个状态，攻防双方可以认为是同时行动，这里的“同时”是一个信息概念而非时间概念，即双方在决策时可能不在同一时刻，但只要双方在决策时不知道对方在当前状态的选择，就认为双方是“同时”行动。

上述分析将每个状态的攻防对抗抽象为一个双人、非零和不完全信息静态博弈。

2) 建立网络状态之间的转移模型

网络状态转移不仅受攻防动作对的影响，还受到很多其他已知和未知因素的影响，转移过程存在随机性，所以以概率方式对其进行描述。又由于下一时间片的网络状态只与当前状态有关，与之前状态无关，即状态的转移符合一阶马尔可夫假设，所以转移概率有 $t(s_{t+1}|s_t, a_{0t}, d_{0t}) = t(s_{t+1}|s_t, a_t, d_t)$ 。

至此，已将网络攻防抽象为一个双人、非零

和的不完全信息随机博弈，随机博弈由贝叶斯静态博弈与马尔可夫决策组成。在此基础上构建不完全信息随机博弈模型 (II-SGM, incomplete information stochastic game model)。

定义 1 不完全信息随机博弈模型是一个九元组 $II-SGM = (N, S, \Theta, P_A, A, t, \pi, R, U)$ 。

① $N = (\text{attacker}, \text{defender})$ 为参与博弈的局中人，分别代表网络攻击者和防御者。

② $S = (s_1, s_2, \dots, s_n)$ 为随机博弈状态集合，每个博弈状态代表一个网络状态。网络状态一般由硬件类型、软件类型、提供的服务、节点连接关系和攻击者权限等属性组成，网络状态的组成属性越多，模型的准确性越好，但复杂度也越高，也更难以分析，需要根据网络选择具体的网络状态表示方式。

③ $\Theta = (\theta_i | i = 1, 2, \dots, n)$ 是攻击者的类型空间。

④ $P_A = (p_A(s_i, \theta_1), p_A(s_i, \theta_2), \dots, p_A(s_i, \theta_n))$ 是防御者对攻击者类型分布的概率判断。

⑤ $A = (A^a, A^d)$ 为攻防动作集合， $A_k^d = \{d_1, d_2, \dots, d_m\}$ 为防御者在状态 S_k 的动作集合，其中， d_i 为防御动作。

⑥ t 为状态转移函数，其中， $t(s'|s, a, d)$ 为在状态 s 攻防双方采取动作对 (a, d) 后，网络转移到状态 s' 的概率， $\sum_{s' \in S} t(s'|s, a, d) = 1$ 。

⑦ $\pi = (\pi_a, \pi_d)$ 为攻防策略， $\pi_a(s_k, \theta_i) = (\sigma_a(s_k, a_1, \theta_i), \dots, \sigma_a(s_k, a_m, \theta_i))$ 为类型 θ_i 的攻击者在网络状态 s_k 的策略， $\sigma_a(s_k, a_m, \theta_i)$ 为其选择动作 a_m 的概率，其中， $\sum_{a_m \in A_k^a} \sigma_a(s_k, a_m, \theta_i) = 1$ 。

⑧ $R = (R_a, R_d)$ 为立即回报，其中， $R_d(s, a, d, \theta_i)$ 表示在状态 s 攻击者类型为 θ_i 时，攻防双方采取动作对 (a, d) 后防御者的立即回报。

⑨ $U = (Q, V)$ 为攻防双方的准则函数集合，其中， Q 为攻防动作准则函数，用于判断攻防动作的优劣， V 为攻防策略准则函数，用于判断攻防策略的优劣。网络攻防是一个序列式决策问题，此类问题常用的准则函数有平均累积收益函数和折扣累积收益函数2种。由于网络攻防对抗中收益与时间有较大关系，所以本文选用折扣累积收益函数作为攻防双方的准则函数。使用折扣因子 $\gamma (0 < \gamma < 1)$ 代表攻防双方对当前回报与未来回报的偏好。具体准则函数如式(1)和式(2)所示。

$$Q_h(s, a, d, \theta_j) = R_h(s, a, d, \theta_j) + \gamma \sum_{s' \in S} t(s' | s, a, d) V_h(s', \pi_a^*(s', \theta_j), \pi_d^*(s'), \theta_j) \quad (1)$$

$$V_h(s_i, \pi_a(s, \theta_j), \pi_d(s), \theta_j) = \sum_{a \in A_i} \sigma_a(s, a, \theta_j) \sum_{d \in D_i} \sigma_d(s, d) Q_h(s, a, d, \theta_j) \quad (2)$$

利用海萨尼公理，通过引入虚拟参与人“自然”，将防御者对攻击收益的不完全信息转换为对“自然”行动的不完美信息，这样就可以使用标准技术对其进行博弈分析。在这些基础上构建一个时间片的博弈树，如图1所示。博弈由上一个时间片转移到当前时间片后，“自然” N 按照转移概率先选择当前时间片的网络状态，再按防御者对攻击者类型分布的概率判断选择攻击者的类型，攻击者 A 和防御者 D 都能观察 N 对状态的选择，但只有攻击者能观察到 N 对攻击者类型的选择。 A 和 D 观察后依据策略选择自己的动作，当前时间片的博弈结束，博弈转移到下一个时间片。

2.3 稳定贝叶斯纳什均衡策略分析及求解

定义2 贝叶斯纳什均衡。在博弈状态 s_i 上，对于所有 θ_j 的所有攻击策略 $\pi_a(s_i, \theta_j)$ 有

$$V_a(s_i, \pi_a^*(s_i, \theta_j), \pi_d^*(s_i), \theta_j) \geq V_a(s_i, \pi_a(s_i, \theta_j), \pi_d^*(s_i), \theta_j)$$

对于所有防御策略 $\pi_d(s_i)$ 有

$$\begin{aligned} & \sum_{j=1}^n p_A(s_i, \theta_j) V_d(s_i, \pi_a^*(s_i, \theta_j), \pi_d^*(s_i), \theta_j) \\ & \geq \sum_{j=1}^n p_A(s_i, \theta_j) V_d(s_i, \pi_a^*(s_i, \theta_j), \pi_d(s_i), \theta_j) \end{aligned}$$

则策略 $(\pi_a^*(s_i, \theta_1), \dots, \pi_a^*(s_i, \theta_n), \pi_d^*(s_i))$ 是网络状态 s_i 上的一个贝叶斯纳什均衡。

若定义2中的收益为马尔可夫稳定收益 $U^* = (Q^*, V^*)$ ，则相应均衡为稳定贝叶斯纳什均衡。

网络攻防对抗具有策略依存性，攻防双方的最优策略依据对手策略的变化而变化，双方在制定策略时，会对对手的策略进行预测，根据预测结果选择自己的最优策略，也会预测到对手对自己的预测，从而进一步调整自己的策略，双方的预测和调整是一个螺旋上升过程，而贝叶斯纳什均衡就是这个螺旋上升过程的最终状态。根据贝叶斯纳什均衡定义可知，均衡策略是攻防双方根据对手策略做出的最优响应，任何一方无法通过单方面努力提高自己的收益。综上可知，利用贝叶斯纳什均衡进行攻击预测和防御策略选取是一种合理、有效的方法。

定理1 不完全信息随机博弈模型 $II-SGM = (N, S, \Theta, P_A, A, t, \pi, R, U)$ 存在稳定贝叶斯纳什均衡。

证明 $II-SGM$ 模型中博弈参与人 N 为攻击者和防御者，是一个有限集。 S 由网络状态组成， Θ 为攻击者类型， A 为攻防动作，三者也都是有限集，故 $II-SGM$ 模型是一个有限不完全信息随机博弈模型，根据文献[14]可知，任何有限不完全信息随机博弈都存在稳定贝叶斯纳什均衡。所以， $II-SGM$ 模型存在稳定贝叶斯纳什均衡解。

Shapley^[15]提出了一种值迭代方法求解由矩阵博弈与马尔可夫决策结合的随机博弈，但是该

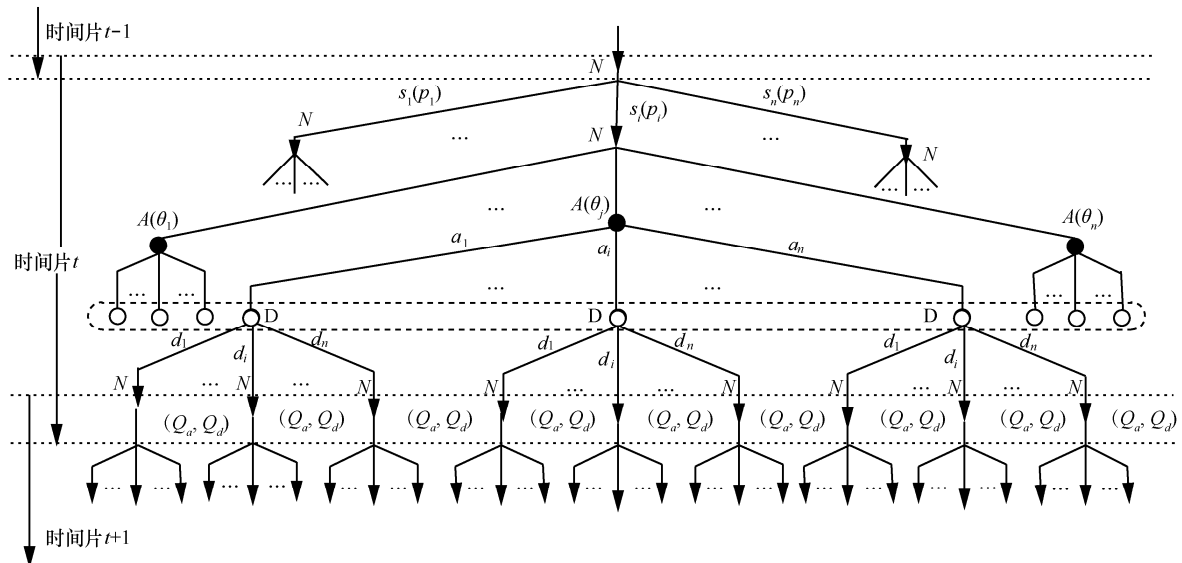


图1 攻防博弈树

方法计算复杂度高, 求解效率低。Filar^[16]等将由矩阵博弈和马尔可夫决策结合的随机博弈的求解问题转化为一个非线性规划问题, 该方法是目前求解随机博弈的主流方法。本文在 Filar 等的基础上给出由静态贝叶斯博弈与马尔可夫决策结合的随机

博弈的求解方法。

定理 2 对于给定不完全信息随机博弈模型 $II-SGM$ 及其收益 V^* 和策略 π^* , V^* 为稳定收益值、 π^* 为稳定贝叶斯纳什均衡解的充要条件为: 收益 V^* 和策略 π^* 为非线性规划 np 的最优解。

$$np = \begin{cases} \min[\sum_{\theta_j \in \Theta} \sum_{s_k \in S} \sum_{h \in N} V_h(s_k, \pi_a(s_k, \theta_j), \pi_d(s_k), \theta_j) - R_h(s_k, \pi_a(s_k, \theta_j), \pi_d(s_k), \theta_j) \\ - \gamma \sum_{s' \in S} t(s' | s_k, \pi_a(s_k, \theta_j), \pi_d(s_k)) V_h(s', \pi_a(s', \theta_j), \pi_d(s'), \theta_j)] \\ \text{s.t.} \\ \textcircled{1} \sum_{d \in A_k^d} \sigma_d(s_k, d) [R_a(s_k, a, d, \theta_j) + \gamma \sum_{s' \in S} t(s' | s_k, a, d) V_d(s', \pi_a(s', \theta_j), \pi_d(s'), \theta_j)] \\ \leq V_a(s_k, \pi_a(s_k, \theta_j), \pi_d(s_k), \theta_j), s_k \in S, \theta_j \in \Theta, d \in A^d \\ \textcircled{2} \sum_{j=1}^n p_A(s_k, \theta_j) \sum_{a \in A_k^a} \sigma_a(s_k, a, \theta_j) [R_d(s_k, a, d, \theta_j) + \gamma \sum_{s' \in S} t(s' | s_k, a, d) V_d(s', \pi_a(s', \theta_j), \pi_d(s'), \theta_j)] \\ \leq \sum_{j=1}^n p_A(s_k, \theta_j) V_d(s_k, \pi_a(s_k, \theta_j), \pi_d(s_k), \theta_j), s_k \in S, a \in A^a \\ \textcircled{3} \sum_{d \in A_k^d} \sigma_d(s_k, d) = 1, \sum_{a \in A_k^a} \sigma_a(s_k, a, \theta_j) = 1, s_k \in S, \theta_j \in \Theta \\ \textcircled{4} \sigma_d(s_k, d) \geq 0, \sigma_a(s_k, a, \theta_j) \geq 0, s_k \in S, \theta_j \in \Theta, d \in A_k^d, a \in A_k^a \end{cases}$$

证明

1) 必要性

若 V^* 为稳定收益值、 π^* 为稳定贝叶斯纳什均衡解, 根据定义 2 可知, π^* 是针对对手策略的最优响应, 故满足约束条件①、②, 约束条件③、④是对自变量 π 取值范围的线性约束, 故 V^* 和 π^* 满足约束条件③、④。由于 V^* 和 π^* 分别是稳定马氏收益和稳定马氏策略, 故

$$\begin{aligned} & V_h^*(s_k, \pi_a^*(s_k, \theta_j), \pi_d^*(s_k), \theta_j) = \\ & R_h(s_k, \pi_a^*(s_k, \theta_j), \pi_d^*(s_k), \theta_j) + \\ & \gamma \sum_{s' \in S} t(s' | s_k, \pi_a^*(s_k, \theta_j), \pi_d^*(s_k)) \\ & V_h^*(s', \pi_a^*(s', \theta_j), \pi_d^*(s'), \theta_j) \end{aligned} \quad (3)$$

由式(3)可知, 当收益和策略取值为 V^* 和 π^* 时, 目标函数为 0。又由约束条件①、②可推出目标函数大于等于 0, 故 V^* 和 π^* 为非线性规划 np 的最优解。

2) 充分性

若 V^* 和 π^* 为非线性规划 np 的最优解, 则式(3)成立, 故 V^* 和 π^* 分别是稳定马氏收益和稳定马氏策略。由约束条件①可知, π_a^* 是攻击者针

对防御者策略的最优响应, 由约束条件②可知, π_d^* 是防御者针对攻击者策略的最优响应。故策略 π^* 是贝叶斯纳什均衡。综上, 可知 V^* 为稳定马氏收益, π^* 为稳定贝叶斯纳什均衡解。

证毕。

3 实例分析

3.1 攻防场景

使用本文方法对一个典型企业网络进行分析, 其拓扑结构如图 2 所示。攻击者来自外网, 远程对内网发起攻击。防御者为内网管理员, 负责全网的网络安全防御。防火墙的策略为外网正常用户只能与内网的 Web 服务器和内网用户通信, 且只能访问 Web 服务器的 HTTP 和 FTP 服务, 其他网络节点和端口均进行阻断。参考 NVD 数据库^[17]给出企业网络脆弱性信息, 如表 1 所示。参考美国麻省理工大学的攻防行为数据库^[18]给出防御动作, 如表 2 所示。

3.2 场景建模

参与人 $N=(attacker, defender)$ 为外网攻击者与内网防御者。

攻防随机博弈状态 $S=(s_0, s_1, s_2, s_3, s_4, s_5, s_6)$ 。

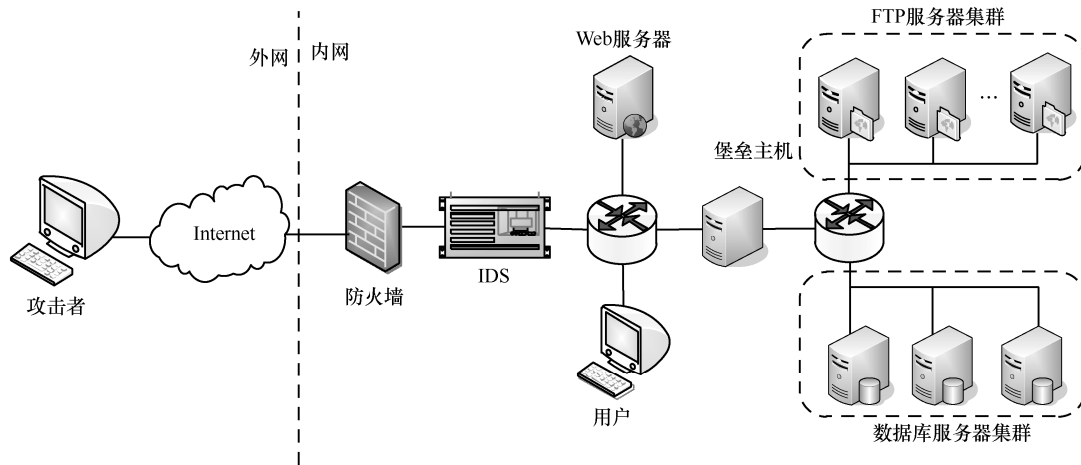


图 2 实验网络拓扑结构

表 1

实验网络脆弱性信息

序号	主机	服务	脆弱性	CVE 编号	目标权限
a_1	Web 服务器	AP	Apache Chunked-Enc	CVE-2015-1635	root
a_2	内网用户	RPC	Code Injection	CVE-2017-7269	root
a_3	堡垒主机	SMTP	Remote buffer overflow	CVE-2014-8517	root
a_4	FTP 服务器	FTP	Write \$home/.rhost	CVE-2014-3556	root
a_5	数据库服务器	MySQL	Local buffer overflow	CVE-2013-4730	user
a_6	数据库服务器	MySQL	Local buffer overflow	CVE-2013-4730	root
a_7	数据库服务器	MySQL	Local buffer overflow	CVE-2016-6662	root

表 2

实验网络防御动作

序号	防御动作
d_1	Limit packets from ports
d_2	Limit SYN/ICMP packets
d_3	Install Oracle patches
d_4	Renew root data
d_5	Uninstall delete Trojan
d_6	Limit access to MDSYS
d_7	Restart Database Server
d_8	Delete suspicious account

为了更清楚地显示网络安全状态的不同，这里只选择具有代表性的主机权限对网络安全状态进行描述。 S_0 ：攻击者在入侵主机具有 root 权限； S_1 ：攻击者在 Web 服务器具有 root 权限； S_2 ：攻击者在内网用户主机具有 root 权限； S_3 ：攻击者在堡垒主机具有 root 权限； S_4 ：攻击者在 FTP 服务器具有 root 权限； S_5 ：攻击者在数据库服务器具有 user

权限； S_6 ：攻击者在数据库服务器具有 root 权限。

攻击者类型 $\Theta = (\theta_1, \theta_2)$ ，其中， θ_1 为高能力攻击者， θ_2 为低能力攻击者。

防御者对攻击者类型的概率判断为

$$(p_A(s_0, \theta_1), p_A(s_0, \theta_2); p_A(s_1, \theta_1), p_A(s_1, \theta_2); p_A(s_2, \theta_1), p_A(s_2, \theta_2); p_A(s_3, \theta_1), p_A(s_3, \theta_2); p_A(s_4, \theta_1), p_A(s_4, \theta_2); p_A(s_5, \theta_1), p_A(s_5, \theta_2); p_A(s_6, \theta_1), p_A(s_6, \theta_2)) = (0.1, 0.9; 0.2, 0.8; 0.15, 0.85; 0.3, 0.7; 0.2, 0.8; 0.4, 0.6; 0.5, 0.5)$$

攻击者动作 $A^a = (A_0^a, A_1^a, A_2^a, A_3^a, A_4^a, A_5^a, A_6^a)$ 。其中， $A_0^a = \{\Phi, a_1, a_2\}$ 、 $A_1^a = \{\Phi, a_3\}$ 、 $A_2^a = \{\Phi, a_3\}$ 、 $A_3^a = \{\Phi, a_4, a_5\}$ 、 $A_4^a = \{\Phi, a_7\}$ 、 $A_5^a = \{\Phi, a_6, a_7\}$ 、 $A_6^a = \{\Phi\}$ 。

防御者动作 $A^d = (A_0^d, A_1^d, A_2^d, A_3^d, A_4^d, A_5^d, A_6^d)$ 其中 $A_0^d = \{\Phi, d_1, d_2\}$ 、 $A_1^d = \{\Phi, d_3, d_4\}$ 、 $A_2^d = \{\Phi, d_3, d_4\}$ 、 $A_3^d = \{\Phi, d_5, d_6\}$ 、 $A_4^d = \{\Phi, d_3, d_4\}$ 、 $A_5^d = \{\Phi, d_7, d_8\}$ 、 $A_6^d = \{d_3, d_4, d_6\}$ 。

转移概率 t 如表 3 所示。

表 3

转移概率 t

S_0				S_1				S_2				S_3					
	S_0	S_1	S_2		S_0	S_1	S_3		S_0	S_2	S_3		S_1	S_2	S_3	S_4	S_5
(Φ, Φ)	1	0	0	(Φ, Φ)	0	1	0	(Φ, Φ)	0	1	0	(Φ, Φ)	0	0	1	0	0
(Φ, d_1)	1	0	0	(Φ, d_3)	0.6	0.4	0	(Φ, d_3)	0.5	0.5	0	(Φ, d_5)	0.4	0.4	0.2	0	0
(Φ, d_2)	1	0	0	(Φ, d_4)	0.9	0.1	0	(Φ, d_4)	0.7	0.3	0	(Φ, d_6)	0.5	0.5	0	0	0
(a_1, Φ)	0.2	0.8	0	(a_3, Φ)	0	0.4	0.6	(a_3, Φ)	0	0.3	0.7	(a_4, Φ)	0	0	0.2	0.8	0
(a_1, d_1)	0.7	0.3	0	(a_3, d_3)	0.5	0.2	0.3	(a_3, d_3)	0.4	0.3	0.3	(a_4, d_5)	0.3	0.3	0.1	0.3	0
(a_1, d_2)	0.9	0.1	0	(a_3, d_4)	0.7	0.1	0.2	(a_3, d_4)	0.6	0.2	0.2	(a_4, d_6)	0.4	0.4	0.1	0.1	0
(a_2, Φ)	0.2	0	0.8									(a_5, Φ)	0	0	0.1	0	0.9
(a_2, d_1)	0.8	0	0.2									(a_5, d_5)	0.3	0.3	0.2	0	0.2
(a_2, d_2)	0.9	0	0.1									(a_5, d_6)	0.4	0.4	0.1	0	0.1
S_4				S_5				S_6									
	S_3	S_4	S_6		S_3	S_5	S_6		S_4	S_5	S_6		S_4	S_5	S_6		S_6
(Φ, Φ)	0	1	0	(Φ, Φ)	0	1	0	(Φ, d_3)	0.4	0.4	0.2						
(Φ, d_3)	0.7	0.3	0	(Φ, d_7)	0.8	0.2	0	(Φ, d_4)	0.5	0.5	0						
(Φ, d_4)	0.9	0.1	0	(Φ, d_8)	0.9	0.1	0	(Φ, d_6)	0.5	0.5	0						
(a_7, Φ)	0	0.4	0.6	(a_6, Φ)	0	0.3	0.7										
(a_7, d_3)	0.4	0.3	0.3	(a_6, d_7)	0.6	0.2	0.2										
(a_7, d_4)	0.6	0.3	0.1	(a_6, d_8)	0.7	0.2	0.1										
				(a_7, Φ)	0	0.1	0.9										
				(a_7, d_7)	0.5	0.2	0.3										
				(a_7, d_8)	0.7	0.1	0.2										

参考文献[9]的方法对立即回报 R 进行量化, 不同类型攻击者的量化结果分别如表 4 和表 5 所示。

表 4

 θ_i 立即回报 R

S_0		S_1		S_2		S_3		S_4		S_5		S_6	
(Φ, Φ)	0,0	(Φ, Φ)	15,-50	(Φ, Φ)	10,-45	(Φ, Φ)	15,-50	(Φ, Φ)	20,-55	(Φ, Φ)	20,-60	(Φ, d_3)	30,-20
(Φ, d_1)	0,-20	(Φ, d_3)	0,-25	(Φ, d_3)	0,-20	(Φ, d_5)	0,-15	(Φ, d_3)	0,-25	(Φ, d_7)	0,-20	(Φ, d_4)	20,-15
(Φ, d_2)	0,-30	(Φ, d_4)	0,-30	(Φ, d_4)	0,-35	(Φ, d_6)	0,-20	(Φ, d_4)	0,-30	(Φ, d_8)	0,-25	(Φ, d_6)	15,-10
(a_1, Φ)	60,-80	(a_3, Φ)	65,-85	(a_3, Φ)	70,-90	(a_4, Φ)	75,-90	(a_7, Φ)	80,-95	(a_6, Φ)	80,-95		
(a_1, d_1)	30,-20	(a_3, d_3)	55,-15	(a_3, d_3)	55,-15	(a_4, d_5)	60,-15	(a_7, d_3)	60,-25	(a_6, d_7)	55,-25		
(a_1, d_2)	20,-10	(a_3, d_4)	30,-5	(a_3, d_4)	25,-10	(a_4, d_6)	30,-15	(a_7, d_4)	30,-15	(a_6, d_8)	25,-15		
(a_2, Φ)	70,-90					(a_5, Φ)	80,-90			(a_7, Φ)	85,-95		
(a_2, d_1)	40,-15					(a_5, d_5)	65,-20			(a_7, d_7)	55,-30		
(a_2, d_2)	25,-10					(a_5, d_6)	35,-15			(a_7, d_8)	15,-10		

表 5

 θ_2 立即回报 R

S_0	S_1	S_2	S_3	S_4	S_5	S_6
(Φ, Φ) 0,0	(Φ, Φ) 10,-40	(Φ, Φ) 5,-35	(Φ, Φ) 10,-55	(Φ, Φ) 15,-35	(Φ, Φ) 10,-55	(Φ, d_3) 20,-20
(Φ, d_1) 0,-35	(Φ, d_3) 0,-35	(Φ, d_3) 0,-25	(Φ, d_5) 0,-25	(Φ, d_3) 0,-40	(Φ, d_7) 0,-25	(Φ, d_4) 15,-15
(Φ, d_2) 0,-25	(Φ, d_4) 0,-30	(Φ, d_4) 0,-45	(Φ, d_6) 0,-25	(Φ, d_4) 0,-35	(Φ, d_8) 0,-30	(Φ, d_6) 15,-5
(a_1, Φ) 50,-60	(a_3, Φ) 45,-75	(a_3, Φ) 65,-80	(a_4, Φ) 60,-80	(a_7, Φ) 60,-75	(a_6, Φ) 65,-75	
(a_1, d_1) 25,-25	(a_3, d_3) 40,-15	(a_3, d_3) 55,-20	(a_4, d_5) 55,-25	(a_7, d_3) 55,-25	(a_6, d_7) 50,-20	
(a_1, d_2) 15,-5	(a_3, d_4) 40,-5	(a_3, d_4) 25,-15	(a_4, d_6) 25,-20	(a_7, d_4) 25,-10	(a_6, d_8) 15,-15	
(a_2, Φ) 55,-80			(a_5, Φ) 65,-75		(a_7, Φ) 60,-60	
(a_2, d_1) 25,-15			(a_5, d_5) 55,-25		(a_7, d_7) 30,-30	
(a_2, d_2) 15,-10			(a_5, d_6) 35,-15		(a_7, d_8) 15,-10	

3.3 稳定贝叶斯均衡求解

根据定理 2, 将 3.2 节中的不完全信息随机博弈模型求解问题转移为非线性规划问题 ($\gamma=0.4$)。使用 Matlab 的求解工具对非线性规划求解, 求解结果如表 6 所示。

3.4 结果分析

由上文的分析可知, 稳定贝叶斯纳什均衡中的攻击策略是对攻击行为的可信预测。根据不同攻击类型的均衡策略及攻击者类型分布概率可以计算出每个状态下脆弱性被利用的概率。

s_0 : [0.3, 0.46, 0.24] s_1 : [0.39, 0.61] s_2 : [0.48, 0.52]
 s_3 : [0.14, 0.41, 0.45] s_4 : [0.58, 0.42] s_5 : [0.32, 0.34, 0.34]
 s_6 : [1]

以 s_0 : [0.3, 0.46, 0.24] 为例, 表明攻击者在状态 s_0 受到现有防御措施的影响, 有 0.3 的概率不发起攻击, 有 0.46 的概率选择 a_1 发起攻击, 有 0.24 的概率选择 a_2 发起攻击。如果防御者加大防御投入, 那么攻击者不发起攻击的概率大于 0.3, 可以

增加网络的安全性, 但这对防御者并不是合理的, 盲目增加防御投入, 虽然增加安全性, 却会降低防御者的收益。另外, 攻击者选择 a_1 的概率明显高于 a_2 , 因而对于防御者来说, 应该更加侧重对脆弱性 a_1 的监测。

将稳定贝叶斯纳什均衡作为防御策略, 以 s_0 为例, 由表 6 可知, 此时防御者需要以 0.1 的概率不执行防御动作, 0.4 的概率选择防御动作 d_1 , 以 0.5 的概率选择防御动作 d_2 。这里的防御策略为一个混合策略。相比纯策略, 混合策略能够带来不确定性, 使攻击者无法预测到防御者的实际行动选择, 增加系统的安全性。

4 结束语

基于博弈论的网络安全研究是现在网络安全领域的主要发展方向之一。为解决网络防御决策问题, 本文将复杂的网络攻防对抗描述为一个随机博弈问题。针对网络安全领域现有随机博弈模

表 6

网络攻防博弈均衡及收益值

网络状态	攻击者 θ_1 攻击策略	攻击者 θ_2 攻击策略	防御策略	攻击者 θ_1 收益	攻击者 θ_2 收益	防御收益
S_0	[0.3, 0.1, 0.6]	[0.3, 0.5, 0.2]	[0.1, 0.4, 0.5]	200.8	165.4	-170
S_1	[0.21, 0.79]	[0.44, 0.56]	[0.12, 0.33, 0.55]	193.9	173.6	-202.3
S_2	[0.17, 0.83]	[0.53, 0.47]	[0.03, 0.48, 0.49]	220.5	183.9	-213.4
S_3	[0, 0.3, 0.7]	[0.2, 0.5, 0.3]	[0, 0.44, 0.56]	167.1	145	-120.8
S_4	[0.4, 0.6]	[0.62, 0.38]	[0, 0.55, 0.45]	189.2	167.6	-182.2
S_5	[0.2, 0.4, 0.4]	[0.4, 0.3, 0.3]	[0, 0.55, 0.45]	234.4	200.7	-222.3
S_6	[1]	[1]	[0, 0, 1]	177.7	145.1	-230.2

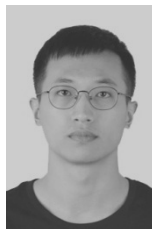
型以完全信息假设为前提,与实际不符的问题,本文将防御者对攻击者收益的不确定性转化为对攻击者类型的不确定性,建立由静态贝叶斯博弈与马尔可夫决策结合的不完全信息随机博弈模型对网络攻防进行分析,通过稳定贝叶斯纳什均衡进行攻击预测和防御策略选取。借鉴现有由矩阵博弈与马尔可夫决策组成的随机博弈的求解方法,给出了本文不完全信息随机博弈模型的非线性规划求解方案。最后通过一个典型的企业网络攻防场景,给出了本文方法的一个详细应用实例。

现有网络状态转移概率往往是由专家知识或历史经验确定,准确性较差。下一步拟将强化学习引入随机博弈中,使防御者通过在线学习修正网络状态转移概率的误差,以获得更高的防御收益。

参考文献:

- [1] LIANG X, XIAO Y. Game theory for network security[J]. IEEE Communications Surveys & Tutorials, 2013, 15(1):472-486.
- [2] WANG Y Z, YU J Y, WEN Q, et al. evolutionary game model and analysis methods for network group behavior[J]. Chinese Journal of Computers, 2015.
- [3] FALLAH M S. A puzzle-based defense strategy against flooding attacks using game theory[J]. IEEE Transactions on Dependable & Secure Computing, 2010, 7(1):5-19.
- [4] 刘玉岭, 冯登国, 吴丽辉, 等. 基于静态贝叶斯博弈的蠕虫攻防策略绩效评估[J]. 软件学报, 2012, 23(3):712-723.
LIU Y L, FENF D G, WU L H, et al. Performance evaluation of worm attack and defense strategies based on static Bayesian game[J]. Journal of Software, 2012, 23(3):712-723.
- [5] GAO X, ZHU Y F. DDoS defense mechanism analysis based on signaling game model[C]// International Conference on Intelligent Human-Machine Systems and Cybernetics. IEEE, 2013:414-417.
- [6] 张恒巍, 李涛. 基于多阶段攻防信号博弈的最优主动防御[J]. 电子学报, 2017, 45(2):431-439.
ZHANG H W, LI T. Optimal active defense based on multi-stage attack-defense signaling game[J]. Acta Electronica Sinica, 2017, 45(2): 431-439.
- [7] SHEN S, LI Y, XU H, et al. Signaling game based strategy of intrusion detection in wireless sensor networks[J]. Computers & Mathematics with Applications, 2011, 62(6):2404-2416.
- [8] ALLAMIGEON X, GAUBERT S, SKOMRA M. Solving generic nonarchimedean semidefinite programs using stochastic game algorithms[C]// The ACM. ACM, 2016:31-38.
- [9] 姜伟, 方滨兴, 田志宏, 等. 基于攻防随机博弈模型的防御策略选取研究[J]. 计算机研究与发展, 2010, 47(10):1714-1723.
JIANG W, FANG B X, TIAN Z H, et al. Research on defense strategies selection based on attack-defense stochastic game model[J]. Journal of Computer Research and Development, 2010, 47(10): 1714-1723.
- [10] 王长春, 程晓航, 朱永文, 等. 计算机网络对抗行动策略的Markov 博弈模型[J]. 系统工程理论与实践, 2014, 34(9): 2402-2410.
WANG C C, CHENG X H, ZHU Y W, et al. A Markov game model of computer network operation[J]. Systems Engineering - Theory & Practice, 2014, 34(9):2402-2410.
- [11] WEI L, SARWAT A I, SAAD W. Risk assessment of coordinated cyber-physical attacks against power grids: A stochastic game approach[C]//Industry Applications Society Meeting. IEEE, 2016:1-7.
- [12] DING K, DEY S, QUEVEDO D E, et al. Stochastic game in remote estimation under DoS attacks[J]. IEEE Control Systems Letters, 2017, 1(1):146-151.
- [13] HARSANYI J C. Games with Incomplete Information Played by "Bayesian" Players, I-III Part I. The Basic Model[M]. Springer Netherlands, 1982:159-182.
- [14] LEI C, ZHANG H Q, WAN L M, et al. Incomplete information Markov game theoretic approach to strategy generation for moving target defense[J]. Computer Communications, 2018, 116:184-199.
- [15] SHAPLEY L S. Stochastic Games[J]. Proceedings of the National Academy of Sciences of the United States of America, 1953, 39(10):1095-1100.
- [16] FILAR J, VRIEZE K. Competitive Markov decision processes[J]. Springer Berlin, 1996, 36(4):343-358.
- [17] ZHANG S, OU X, CARAGEA D. Predicting cyber risks through national vulnerability database[J]. Information Systems Security, 2015, 24(4-6):194-206.
- [18] GORDON L, LOEB M, LUCYSHYN W, et al. 2015 CSI/FBI computer crime and security survey[C]. Proceedings of the 2014 Computer Security Institute//San Francisco: IEEE, 2015. 48-64.

[作者简介]



杨峻楠(1993-), 男, 河北藁城人, 信息工程大学硕士生, 主要研究方向为网络信息安全、博弈论和强化学习。



张红旗(1962-), 男, 河北遵化人, 博士, 信息工程大学教授、博士生导师, 主要研究方向为网络安全、风险评估、等级保护和信息安全管理。