文章编号: 1006-2467(2008)02-0198-04

基于动态博弈理论的分布式拒绝服务攻击防御方法

张少俊, 李建华, 陈秀真, 胡 威
(上海交通大学 信息安全工程学院, 上海 200240)

摘 要: 将分布式拒绝服务攻击视作一种可观察行动的多阶段不完全信息博弈, 给出了该博弈的扩展型表示, 提出为了达到该博弈的完美贝叶斯均衡, 需解决局中人类型的信念计算及修正问题. 作为观点的实践, 提出一种根据访问速率以及访问流量源地址分布特征对分布式拒绝服务攻击流量进行选择性过滤的方法, 并对该方法进行了验证.

关键词: 网络攻防; 分布式拒绝服务攻击; 博弈论; 防御; 完美贝叶斯均衡
中图分类号: TP 393 文献标识码: A

Method Research for Defending Against Distributed Denial-of-Service Attacks Based on Dynamic Game Theory

ZHANG Shao-jun, LI Jian-hua, CHEN Xiu-zhen, HU Wei
(School of Information Security Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: As one of the most notorious network attacks, distributed denial-of-service (DDoS) attack is famous for its easiness to launch and difficulty to defend. DDoS attack was regarded as a multistage game of incomplete information with observable actions and its extensive form was given out. It is pointed out that to attain the game's perfect Bayes equilibrium, the problem of computing and modifying the belief value of each player's type must be solved. As a practice of this viewpoint, a method of filtering attack stream according to its package rate and source address distribution characteristic was proposed. Finally, simulation was given out to demonstrate the effectiveness of the method.

Key words: network attack and defense; distributed denial-of-service (DDoS) attack; game theory; defending; perfect Bayes equilibrium

随着网络技术的发展, 远程攻击手段层出不穷, 分布式拒绝服务(distributed denial-of- service, DDoS)攻击无疑是其中的佼佼者. 由于其发起简单、破坏性大的特点, DDoS 被世界范围内的攻击者们广泛使用. 我国遭受 DDoS 攻击的情况同样十分普遍, 许多 ISP、ICP 都曾遭到过该类攻击. 尤其是宽带业务的推广和宽带用户数量大幅增加, 使 DDoS 攻击变得更为严重.

随着攻击威胁的增长, 近年来关于如何防御 DDoS 攻击的研究取得了令人振奋的进展. 按照部署位置, DDoS 防御机制一般可以分为源网络防御、中间网络防御和目标网络防御[1]. 目前较易于实施的方法是在可能遭受攻击的目标网络中部署防御措施, 包括更改主机默认设置[2], 根据源地址熵值对请

收稿日期: 2007-03-08
基金项目: 国家自然科学基金(60605019, 60772098, 60672068)资助项目, 教育部新世纪优秀人才支持计划(NCET-06-0393)资助项目
作者简介: 张少俊(1978-), 男, 上海市人, 博士生, 从事计算机网络安全研究, E-mail: zshaojun@sjtu.edu.cn.
李建华(联系人), 男, 教授, 博士生导师, 电话(Tel.): 021-34204532; E-mail: lijh888@sjtu.edu.cn.

求进行动态过滤[3] 等. 这些方法虽能在一定程度上抑制 DDoS 攻击, 但并未对攻击参量进行动态跟踪估计, 阶段调整防御参量, 因而整体防御效果无法达到最优. 本文试图在目标网络中构建一种通用的基于动态博弈理论的 DDoS 防御方法, 动态跟踪估计攻击参量. 通过不断深化对攻击者的认识, 在博弈的每个阶段选择最优策略, 从而实现对 DDoS 攻击流量的有效抑制.

1　动态不完全信息博弈

博弈论与网络攻防行为有着天然的密切联系. 博弈论主要研究理性的组织、团队或者个人在一定的规则下, 同时或先后、一次或多次, 从各自允许选择的行为或策略中进行选择并加以实施, 最终各自取得相应结果的过程.

本质上, 网络攻防是一种动态不完全信息博弈[4], 特别地, 对于包含多个步骤、有一定持续期的网络攻防事件(如 DDoS 攻防), 可将其看作可观察行动的不完全信息多阶段博弈.

多阶段是指博弈可以分成若干周期进行. 例如: DDoS 攻击通常持续几分钟到几昼夜不等. 在这个过程中, 攻击者可以动态地调整攻击方法与攻击参量(如发包速率、源地址伪造的随机特性等), 防御者则可以动态地调整防御参量(如设置不同的防火墙过滤规则)甚至更换防御设备. 不同的攻防方法或参量组合构成了不同的博弈阶段.

不完全信息是指一般情况下, 攻击方对防御方因拒绝服务所导致的损失以及防御方采用的防御方法、参量等信息并不清楚. 同样, 防御方对攻击方的信息也不确定.

可观察行动指攻击方可以通过模拟正常的客户考察攻击效果, 防御方也可以通过模拟正常客户甚至直接监测、分析网络流量考察防御的效果, 并决定后续行动.

在不完全信息动态博弈中, 贝叶斯均衡占有重要的地位, 它是完全信息博弈纳什均衡在不完全信息博弈中的推广. 在该均衡点, 所有局中人的策略选择达到了最优. 与完全信息动态博弈不同的是, 不完全信息动态博弈引入了类型信念函数的概念, 要求任何一个局中人必须依靠博弈历史 h_t 建立对手类型分布的后验信念, 并根据该信念在后续博弈中选择使收益最大化的策略.

2　DDoS 攻防博弈建模

建模中考虑 DDoS 所有可能的外部环境或者攻防过程中所有的不确定因素是困难的. 因此, 本文对 DDoS 攻防过程作如下限定, 以得到一个适当简化但能体现 DDoS 攻防博弈本质的模型:

(1) 博弈包括 3 个局中人: 攻击者 A, 防御者 D, 合法用户 U. U 发送合法的请求数据包, 随后等待 D 的响应. 如 D 对 U 的请求进行了响应, 双方都有一定的赢利, 如果 D 没有响应 U, 则双方没有赢利. A 的目标是破坏这个过程, 通过大量发送伪造地址的请求数据包, 占用 D 的服务资源, 使得 D 对 U 的响应率下降. 假设 3 个参与者都是理性的, 以追求各自盈利最大化为目标.

(2) 不对 DDoS 攻击的具体类型(如 Syn Flood, UDP Flood 等)作假设, 因而不考虑特效的 DDoS 防御手段. 规定 D 可以且仅可以对每一个服务请求选择响应或者不响应.

(3) D 同时只能对一个用户服务, 并且 D 在服务完成后能够判断这次服务的对象是合法用户还是攻击者, 即知道盈利的变动.

如前所述, DDoS 攻防本质上是一种可观察行动的不完全信息多阶段博弈过程. 在攻击的初始时刻, A 选择参数类型 θ_A, U 选择参数类型 θ_U. 对此 D 是无知的, 只能通过博弈历史对其进行估计. 博弈的一个阶段定义为: 首先由 A 或 U 发送一个服务请求, 而后 D 决定是否对该请求进行响应.

博弈一个阶段的扩展型如图 1 所示.

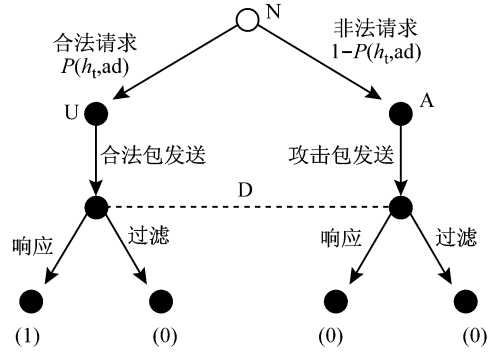

图 1　DDoS 攻防博弈的一阶段

Fig. 1　One stage of DDoS game

图中: N 为虚拟局中人; ad 为服务请求源地址. 图 1 通过海萨尼转换将该不完全信息博弈转化为完全但不完美信息动态博弈. 在博弈初始阶段, N 选择该请求为合法请求的概率为 $P(h_t, \mathrm{ad})$. 该值可认为是 D 对于来自 ad 的服务请求的合法性的信念.

设 D 每次成功地向 U 提供服务可以得到的盈利 $\Delta_{UD}=1$, 服务平均持续时间为 τ, 平均盈利率 δ, 请求平均到达率为 R, 那么, 如果选择响应该请求, 将使 D 在未来 τ 内期望获得盈利:

$\Delta u_{D,1} = P(h_t, \mathrm{ad}) \cdot 1 + (1 - P(h_t, \mathrm{ad})) \cdot 0$

如果选择过滤该请求，D 在 τ 内期望盈利：

$$\Delta u_{D,2} = 0 + \delta_d\left(\tau - \frac{1}{R}\right)$$

为达到盈利最大化，当 $\Delta u_{D,1} > \Delta u_{D,2}$ 时，D 将选择响应请求，此时解得 $P(h_t, \mathrm{ad}) > \delta_d\left(\tau - \frac{1}{R}\right)$. 反之，当 $P(h_t, \mathrm{ad}) < \delta_d\left(\tau - \frac{1}{R}\right)$ 时，D 将选择过滤.

根据博弈的可观察性，δ_d、R、τ 理论上是可以获得的，因而任意一阶段促使 D 决定采取响应策略还是过滤策略的信念阈值是确定的. 因此，只要能够确定 $P(h_t, \mathrm{ad})$，D 就知道该采取什么行动.

3 基于特定类型空间的防御算法

本文假设 A 具有类型空间 $\Theta_A = \{R_A, P_A(\mathrm{ad})\}$. U 具有类型空间 $\Theta_U = \{R_U, P_U(\mathrm{ad})\}$. 其中：$R_A$、$R_U$ 分别为 A 和 U 的请求速率；$P_A(\mathrm{ad})$、$P_U(\mathrm{ad})$分别为 A 和 U 的请求源地址概率分布. 记 $R_A(h_t)$、$R_U(h_t)$、$P_A(h_t, \mathrm{ad})$、$P_U(h_t, \mathrm{ad})$分别为博弈历史 h_t 下 D 对 R_A、R_U、$P_A(\mathrm{ad})$、$P_U(\mathrm{ad})$的估计. 根据贝叶斯公式，容易得到：

$$P(h_t, \mathrm{ad}) = \frac{R_U(h_t) P_U(h_t, \mathrm{ad})}{R_U(h_t) P_U(h_t, \mathrm{ad}) + R_A(h_t) P_A(h_t, \mathrm{ad})}$$

对于 $R_A(h_t)$、$R_U(h_t)$，在请求速率变化不是非常剧烈的情况的下，可以通过统计请求到达前 Δt 时间段内收到的请求的个数 $N_A(t)$和 $N_U(t)$，即：

$$R_A(h_t) \approx \frac{N_A(t)}{\Delta t}$$

$$R_U(h_t) \approx \frac{N_U(t)}{\Delta t}$$

对于 $P_A(h_t, \mathrm{ad})$、$P_U(h_t, \mathrm{ad})$，本文采用识别请求包的源地址特性的方法，近似构造出其概率分布函数. 识别源地址分布特性的方法有许多种，本文采用 PATRICIA 树[5]结合源地址熵计算[3].

对于合法用户，按如下方法构建 $P_U(h_t, \mathrm{ad})$：

(1) 未发生攻击时，取 Δt 为较大值，如 1 h，记录所有源地址出现概率 $P_n(\mathrm{ad})$.

(2) 取 α 为(0, 1] 区间内的常数，对于每一个源地址 ad，定义 $P'_U(h_t, \mathrm{ad}) = \mathrm{e}^{P_n(\mathrm{ad})/\alpha}$.

(3) 归一化. $P_U(h_t, \mathrm{ad}) = \frac{P'_U(h_t, \mathrm{ad})}{\sum_{\mathrm{ad}} P'_U(h_t, \mathrm{ad})}$.

对于攻击者，按如下方法构建 $P_A(h_t, \mathrm{ad})$：

(1) 取 Δt 为较小值，计算主机节点源地址出现概率 $P_n(\mathrm{ad})$和子网节点 sn 熵值 $S(\mathrm{sn})$.

(2) 设定阈值 β，$\beta > 0$. 对于子网节点 sn，若有 $S(\mathrm{sn}) > \beta$，说明该子网内发现有较强随机性的攻击流量，此时向下追溯，直至子网节点 sn'，有 $S(\mathrm{sn}') > \beta$，且 sn'的子节点 sn''都是主机节点或 $S(\mathrm{sn}'') < \beta$. 此时，认为子网 sn'内存在子网源地址伪造. 记$N(\mathrm{sn}')$为 sn'中有效地址的个数，则：

$$P_A(h_t, \mathrm{ad}) = \frac{\sum_{\mathrm{ad}} P_n(\mathrm{ad})}{N(\mathrm{sn}')},\ \mathrm{ad} \in \mathrm{sn}'$$

(3) 对于其他子网，定义 $P_A(h_t, \mathrm{ad}) = P_n(\mathrm{ad})$.

4 算法验证

为了验证算法，本文构建了包含 3 台主机的局域网：主机 1 运行服务；主机 2 模拟合法用户从 172.16.0.0/16 对服务进行访问；主机 3 模拟攻击者，以：①固定源地址 172.16.1.1 ②172.16.0.0/16 内随机选择的源地址 ③172.16.1.0/24 内随机选择的源地址，对服务进行攻击. 如图 2 所示.

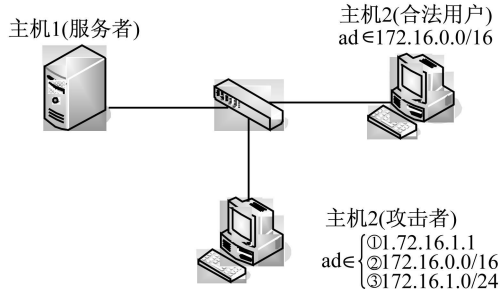

图 2 DDoS 攻防模拟环境

Fig. 2 Simulation environment of DDoS

取 $\tau = 5$ s，合法用户请求速率 $R_U = 5$ 次/s，$\alpha = 0.2$，$\beta = 8.0$，得到盈利曲线如图 3 所示.

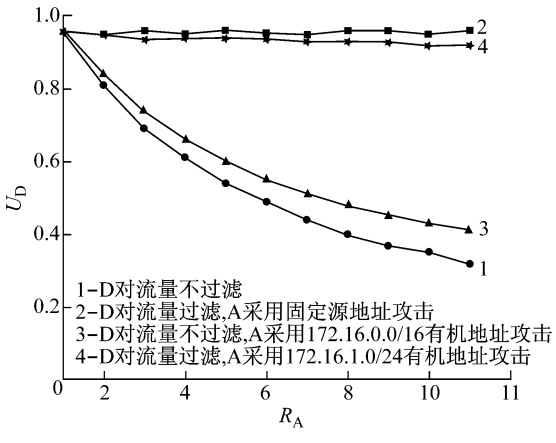

图 3 防御者盈利曲线

Fig. 3 Payoff curves of the defender

由图 3 可见，曲线 1 说明当攻击者加大攻击流量时，防御者的盈利有显著下降；曲线 2 说明算法可

以很好地识别来自固定源地址的攻击流量, 无论攻击者如何加大流量, 盈利始终没有明显的下降; 曲线 4 基本与曲线 2 重合, 说明当攻击者采用 172.16.1.0/24内的随机源地址进行攻击时, 攻击流量与合法流量的源地址虽有一定重叠, 但重叠程度相对较小, 盈利曲线虽有小幅下降但仍接近理想盈利. 曲线 3 与曲线 1 较为接近, 这时攻击流量与合法流量都来自子网 172.16.0.0/16, 但两者源地址概率分布由于生成算法不同而略有错开, 因而采取过滤措施后的盈利稍高于不采取过滤措施时的盈利.

4　结　语

本文对 DDoS 攻防过程进行了博弈论建模, 指出了信念函数在攻防博弈中的重要地位. 作为该观点的实践, 在合理假设博弈各方类型空间的基础上提出了一种流量选择性过滤算法. 在本文提出的博弈模型中, 对 DDoS 攻防过程设置了若干假设. 在现实环境中面对各类 DDoS 攻击时, 这些假设并非都能得到满足. 作为下一步研究, 计划对 DDoS 攻防进行更为准确和贴近实际的建模, 同时扩展局中人的类型空间, 以使局中人信念的表达更为全面.

参考文献:

[1] Mirkovic J, Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms[J]. **Computer Communication Review**, 2004, 34(2): 39—53.

[2] Min F, Zhang J, Li W. Tradeoffs of DDoS solutions [C]// **Parallel and Proceedings of the 4th International Conference on Distributed Computing, Applications and Technologies**. Chengdu: IEEE, 2003: 198—200.

[3] Feinstein L, Schnackenberg D, Balupari R, *et al*. Statistical approaches to DDoS attack detection and response[C] // **Proc of the DARPA Information Survivability Conf and Exposition**. Washington: IEEE, 2003: 303—314.

[4] Xia Z, Zhang S. A kind of network security behavior model based on game theory[C] // **Proceedings of the 4th International Conference on Parallel and Distributed Computing, Applications and Technologies**. Chengdu: IEEE, 2003: 950—954.

[5] Morrision D R. PATRICIA-practical algorithm to retrieve information coded in alphanumeric[J]. **Journal of ACM**, 1968, 15(4): 514—534.

(上接第 **197** 页)

[6] Sun Qing-xian, Fang Tao, Guo Da-zhi. Study on scale transformation in spatial data mining[C]// TANG Xin-ming. **Proceedings of International Symposium on Spatio-temporal Modeling, Spatial Reasoning, Spatial Analysis, Data Mining and Data Fusion**. Beijing: Peking University Press, 2005: 275—279.

[7] Wang Shu-liang, Li De-ren. A perspective of spatial data mining[C]//Gong Jian-ya. **Geospatial Information, Data Mining, and Applications**. Wuhan: Wuhan University Press, 2005: 604518-1—604518-10.

[8] 行小帅, 焦李成. 数据挖掘的聚类方法[J]. 电路与系统学报, 2003, 8(1): 59—67.
XING Xiao-shuai, JIAO Li-cheng. Clustering method in the field of data Mining[J]. **Journal of Circuits and Systems**, 2003, 8(1): 59—67.

[9] Albayrak S, Amasyal F. Fuzzy C-means clustering on medical diagnostic systems[DB/OL] (2003-05-06) [2006-12-07]. http://www.ce.yildiz.edu.tr/mygetfile.php?id=269.

[10] 汤孝琴, 戴汝源. 数据挖掘中聚类分析的技术方法[J]. 微计算机信息, 2003, 19(1): 3—4.
TANG Xiao-qin, DAI Ru-yuan. Technique of cluster analysis in data mining[J]. **Control & Automation**, 2003, 19(1): 3—4.

[11] 王美华. 数据挖掘领域中的聚类方法[J]. 南华大学学报(理工版), 2004, 18(1): 58—62.
WANG Mei-hua. Clustering algorithm in data mining [J]. **Journal of Nanhua University (Science & Engineering Edition)**, 2004, 18(1): 58—62.