网络敌手模型研究

蒋建春[1]，文伟平[2]
（1.中国科学院软件研究所，2.北京大学）

【摘要】本文首先分析网络攻击特性，给出其形式定义、研究分析了网络攻击特点及攻击模型发展状况，提出了一个网络敌手模型，该模型由三个子模型组成，即网络敌手心智子模型、网络敌手攻击决策子模型、网络敌手攻击行为变迁子模型。这三个子模型将刻画网络攻击敌手的特征，包括敌手内部心智及外在的行为。该模型克服已有模型的不足，模型的能力也更强，能够描述攻击者的意图、决策及行为。并进行形式化分析网络敌手的特征、决策规划、攻击行为过程等网络攻击环节。其目的在于知己知彼，为网络信息安全对抗技术及网络安全策略的研究提供理论基础和指导依据。
【关键词】敌手 攻击 模型 网络

相关研究工作

网络敌手模型的研究有利于制定对应的安全策略及采取合适的安全措施，因而有关攻击模型的工作受到研究人员的重视。目前模型研究工作可以分成三类，下面分别论述：

1．第一类是基于图形模型；攻击活动以图形方式刻画。最简单易理解模型就是 Schneier 所提出攻击树（attack tree），树的根及节点对应攻击者总目标和子目标，而叶子表示攻击操作，树中的节点和叶子的关系有“与”和“或”。McDermott 等人在进行安全渗透实验中，以 Petri 网来描述攻击活动，简称攻击网（attack net），攻击网模型将状态（如控制或知识）及安全相关的系统或实体表示成 Petri 网的库所（place），类似攻击树的节点，而 Petri 网的迁移（Transition）表示攻击者动作（attacker action），有向弧线就把库所和迁移连接起来，网中托肯 (token) 的移动就表示攻击变化过程；Sheyner 等人以攻击图 AG（attack graph）来描述攻击活动，即：

$$AG = \langle S, \tau, S_0, S_s \rangle$$

其中 S 是状态集合，$\tau \subseteq S \times S$ 是迁移函数，$S_0 \subseteq S$ 是初始状态，$S_s \subseteq S$ 是攻击者的目标状态集合。攻击者的执行操作片断 (execution fragment) 定义成有限状态序列 $s_0 s_1 \cdots s_n$，其中要求 $(s_i, s_{i+1}) \in \tau, 0 \le i \le n$，若攻击者的执行操作片断中的最后状态包含于 S，则攻击者的操作就是一个攻击，对应于攻击者的原子攻击，所有这些执行片断就形成了攻击过程。

2．第二类是基于语言描述型模型。攻击语言（attack language）的研究受到广泛关注，现在有关描述攻击活动语言共有六种基本类型：事件语言（Event languages）、响应语言（Response languages）、报告语言（reporting languages）、关联语言（Correlation languages）、挖掘语言 (Exploit languages)、检测语言（Detection languages）。这些语言从不同角度描述攻击不同特征。攻击语言从某个角度刻画了网络攻击的局部情况，例如攻击的特征，但难以描述攻击动态情况、攻击决策过程以及网络攻击者的情况。

3．第三类是基于攻击过程模型。Erland 将攻击过程分成学习阶（learning phrase）、标准阶段（standard attack phrase）和创造性攻击阶段（innovative attack phrase）三个阶段，该模型只能从刻画攻击者能力的变迁，而未能就网络攻击整体情况进行描述。

（Cyber-Terrorist）的行为过程模型，这个模型从策略、拥有资源、具备智力、风险接受度、特定攻击目标、攻击过程等角度来刻画赛博恐怖分子模型。赛博恐怖分子攻击过程模型中由情报收集、计划准备、目标网络发现、测试实验、风险判断、攻击执行、破坏效果评估组成。但是侧重于描述攻击决策，而就攻击操作行为执行实际情况不能体现出来。虽然赛博恐怖分子是假设存在的，目前还未确定，美国 DARPA 投入经费进行研究。Bradley J. Wood 提出内部威胁模型（Insider Threat Model），此模型描述内部威胁者属性，包括访问（ACCESS）、知识（Knowledge）、特权（Privilege）、技巧（Skill）、风险（Risk）、策略（Tactics）、动机 (Motivation) 和步骤 (Process)，目的在于仿真内部网中敌手，该模型尚处于发展初期，当前还没有一个完全成熟模型。兰德公司在一份研究报告中给出内部威胁者粗略模型，认为模型有人、工具、环境三部分，包含四个基本元素：观测元素，用于模型测量；轮廓元素，定义人、工具、环境的框架或结构；行为元素，定义特征、属性、关系；F（X），定义模型的功能。

此模型的开发意图就是用于预测、检测、响应、报警、策略开发、教育培训等。美国的 DARPA 的系统保障方法采用了 CARO 模型，此模型从目标 (Objectives)、风险（Risk）、能

力（Capability）、访问（Access）等角度来描述网络敌手的特征；而 James K. Williams 通过调换 CARO 模型的顺序，提出 ORCA 网络敌手模型，ORCA 模型将分成几个阶段。

综合上述，网络攻击模型的研究尚处于研究发展阶段，目前还没有一个世界上公认模型，已有的网络攻击模型存在不足之处，主要缺陷有：模型描述能力的局限性，例如现有模型中不能够刻画网络攻击敌手的心智活动；模型描述的非完整性，模型受限于网络敌手的某些有限行为，无法刻画网络敌手攻击的全部过程；模型描述的粗粒度，模型无法刻画攻击敌手的更细粒度操作行为变迁。

网络敌手模型

1．基本概念定义与符号

定义1　网络敌手

指攻击网络的成员或组织，包括个体、组织、国家等，例如黑客团体。以小写字母符号 s 表示，大写字母 S 表示敌手集合。

定义2　攻击意图

网络敌手所希望攻击实现的目的，如经济利益、报复等。以符号 i 表示，大写字母表示 I 意图集。

定义3　攻击对象

攻击者的操作对象，例如文件、主机或路由器。以符号 o 表示，大写字母 O 表示攻击对象集。

定义4　网络攻击操作

网络敌手对攻击对象施加的动作，以符号 a 表示，大写字母 A 表示攻击动作集。

定义5　攻击目标

攻击者实现攻击意图时所期望东西。例如攻击者获取口令文件。以符号 g 表示单个攻击目标，大写字母 G 表示攻击目标集。

定义6　网络敌手外部环境

网络敌手所处的攻击环境是动态变化，敌手将根据从外部环境所获得的信息来做出有关的判断或行动。本文将网络敌手外部环境抽象成一组消息队列，以符号 M 表示消息集合，而 m 表示消息集中的元素。

定义7　网络攻击敌手信念

网络攻击敌手的信念是攻击者所具有的关于目标网络的信息，它们可能是不完整的，甚至可能是不正确的。这些信息可以分为客观事实(知识)和主观态度，前者的正确性是确定的，如“目标网络提供 WEB 服务”，而后者的正确性是不确定的，如“我相信能够非授权进入 WEB 服务器”。以符号 B 表示网络攻击敌手信念集，b 表示信念集中的元素。

定义8　网络攻击敌手愿望

用于描述网络攻击敌手对未知网络系统状况以及可能采取的行为路线的喜好，属于思维状态的感情方面。敌手的重要愿望特性之一就是拥有互不相容的愿望，而敌手也不必相信他的愿望能否实现。以符号 D 表示网络攻击敌手愿望集，d 表示愿望集中的元素。

定义9　网络攻击敌手知识

网络攻击敌手实现攻击意图前提条件是拥有相应的知识，比如目标网络的通信协议漏洞，或系统操作等。敌手掌握知识多少就决定其是否能够实现他的攻击目标。但攻击敌手的知识是可以改变的，例如某攻击敌手需要获取系统漏洞利用，他可以通过自己学习或从网上搜集。敌手知识也是动态的演变过程。以符号 K 表示网络攻击敌手知识集，k 表示知识集中的元素。

定义10　网络攻击代价

有关攻击过程中的操作开销，例如攻击敌手的计算资源，以符号 c 表示攻击代价，大写字母 C 表示攻击代价集。

2．网络敌手模型概况

网络敌手心智模型、网络敌手攻击决策模型、网络敌手行为变迁模型组成了综合网络敌手模型。各子模型不是相互独立，而是相互关联，相互作用，形成一个有机的综合网络敌手模型。

3．网络敌手心智模型

网络攻击敌手具有多种心智态度，例如愿望、信念、意图、知识、情感、兴趣等。表 2 是网络敌手心智状态变量部分实例表。敌手类型不同，其心智状态不相同，如黑客与信息战士是一样，黑客的意图是尝试安全技术，而信息战士就为攻击特定敌手的目标，获取信息优势；有组织性网络敌手的信念强度高于非组织性的网络敌手。

引用人工智能中 BDI 模型，本文中形式化定义网络攻击敌手心智模型：$MS=\langle G, T, B, D, I, K, M, R\rangle$

其中：

$G: G\times T\times M\rightarrow G$，攻击敌手的目标是时间、外部环境的事件函数，网络敌手目标将随着时间及外部环境的事件而动态变化，生成攻击目标序列；

T 是时间点集合，每个时间点对应于网络敌手的某个心智状态；

$B: B\times T\times M\rightarrow B$，攻击敌手的信念是时间、外部环境的事件函数，网络敌手信念将随着时间及外部环境的事件而动态变化，生成敌手信念序列；

$D: D\times T\times M\rightarrow D$，攻击敌手的愿望是时间函数、外部环境的事件函数，网络敌手愿望将随着时间及外部环境的事件而

动态变化，生成敌手愿望序列；

$I: I\times T\times M\rightarrow I$，攻击敌手的意图是时间函数、外部环境的事件函数，网络敌手意图将随着时间及外部环境的事件而动态变化，生成敌手意图序列；

$K: K\times T\times M\rightarrow K$，攻击敌手的知识是时间函数、外部环境的事件函数，网络敌手知识将随着时间及外部环境的事件而动态变化，生成敌手知识序列；

R 是 G、B、D、I、K 上约束规则集，即敌手的愿望、信念、意图、知识、目标的存在约束关系，如愿望集合包含意图集、敌手的愿望与知识相匹配、目标符合敌手的意图、信念支持目标的实现等。

网络敌手攻击过程中，其心智状态不是保持静止，而是随着攻击进展，心智状态发生演变，如攻击受到追踪，放弃攻击目标。设元$\langle MS^0, next_act, next_ms\rangle$组表示敌手心智状态变化模式，其中 MS^0 是攻击敌手的初始心智状态，由目标、信念、愿望、意图、知识的集合组成；$next\_act: MS\times ACT\rightarrow ACT$，是从心智状态集 MS 到攻击原子操作集 ACT 的选择函数；$next\_ms: MS\times ACT\rightarrow MS$，是指攻击敌手执行一个攻击后，产生新的心智状态的函数。则网络敌手心智状态的演变流程步骤如下：

$step1: MS^0 = InitMetalState(), n=0;$

$step2: a=next\_act(MS^n);$

$step3: Do(a);$

$step4: MS^{n+1}=next\_ms(MS^n,a), n=n+1;$

$step5: goto \quad step2.$

4. 网络敌手攻击决策模型

网络敌手攻击决策不是凭空产生，而是依据自己的心智状态及所掌握的资源、攻击成本代价制定出合适的攻击规划。网络敌手制订规划的目的在于寻找一条攻击操作链。网络攻击操作链由攻击者触发，攻击者应用一定的攻击工具（包括攻击策略与方法），对目标网络和系统进行（合法和非法的）访问，达到一定的攻击效果，实现攻击者预定义的攻击目标，最终攻击敌手的意图。

所有网络攻击链就形成网络攻击规划库集合。将攻击敌手决策规划模型定义成一个四元组：

$attack\_plan=\langle plan\_goal, plan\_premise, plan\_body, plan\_result\rangle$

其中，$plan\_goal$ 表示攻击目标描述；$plan\_body$ 表示攻击规划体，用于描述执行的攻击动作序列或攻击脚本；$plan\_premis$ 表示规划前提，攻击规划体运行必须满足的前提条件集合，即攻击资源；$plan\_result$ 表示规划执行预期结果，在规划前提已满足条件下，规划体执行后可能产生的新的敌手状态集合及攻击效果。下面给出网络敌手攻击规划的BNF语法结构，其中“$\langle a\rangle$+”代表项 a 的一次或多次出现，“；”，“||”，和“*”分别是顺序、并行和迭代型符合动作合成符号，$goal\_list$ 表示目标描述表，$cost\_list$ 表示代价描述表，$resourse\_list$ 表示资源描述表，$capability\_list$ 表示能力描述表，$attack\_list$ 表示网络攻击操作描述表，$result\_list$ 表示网络攻击操作结果描述表。$event\_list$ 表示异常事件描述表。

5. 网络敌手攻击行为变迁模型

攻击行为是网络敌手实现其意图的外在攻击动作表现，按攻击行为的效果可以分成若干类型，下面分别论述。

1）攻击源隐藏行为

隐藏网络攻击敌手主机位置使得系统管理无法追踪。如攻击敌手使用IP地址欺骗。

2）信息收集行为

确定攻击目标并收集目标系统的有关信息。如收集主机帐号信息。

3）弱点挖掘行为

从收集到的目标信息中提取可使用的漏洞信息。如破解目标系统的弱口令。

4）目标权限访问获取行为

获取目标系统的普通或特权帐户的访问权限。如目标的网络访问。

5）攻击操作隐蔽行为

网络攻击敌手为防止攻击操作被发现，而对目标系统进行有关操作。如隐藏攻击进程。

6）目标威胁实施行为

网络攻击敌手按其攻击意图，对目标系统进行破坏活动或者以目标系统为跳板向其他系统发起新的攻击。如拒绝服务攻击。

7）设置远程控制后门行为

此行为目的在于是通过在目标系统中开辟后门，方便以后入侵。如在主机中安装后门工具，开放远程访问端口服务。

8）攻击痕迹清除行为

该行为目的是避免安全管理员的发现、追踪以及法律部门取证。如日志删除。

9）评估攻击效果行为

该行为用于验证是否实现网络敌手的攻击意图。如远程测试WEB网站主页信息。

网络敌手行为是有目的，这就是围绕攻击敌手的意图而进行。为了达到目的，网络敌手就要将不同类型的攻击行为按照预先制定攻击规划而适当地选择和组织。网络敌手攻击行为输出过程用一个五元组$BM=\langle Q, \Sigma, \delta, q_0, F\rangle$来自动机表示，其中：

Q是一个有穷的状态集合，所有的攻击操作集合$Q=A$；

Σ是一个有穷输入字母表，$\Sigma=MS\times PS$，即Σ的元素是由网络攻击敌手心智集合与攻击规划集合的笛卡尔乘积构成；

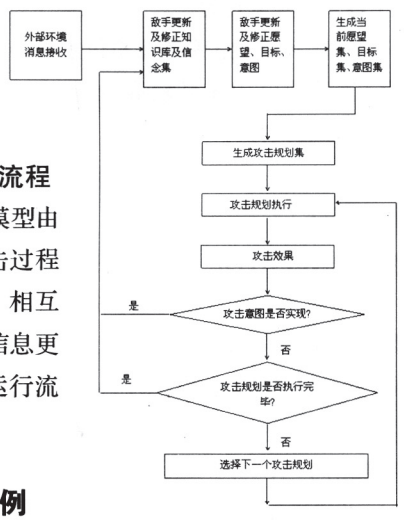
q_0在Q中，它是初始状态，攻击敌手初始操作集；

F是终结状态集合，即网络敌手最终攻击行为集合；

δ是转移函数，$Q\times\Sigma\rightarrow Q$。即输出网络敌手的行为函数。

网络敌手攻击行为不是无关联，而是根据目标系统实际情况及攻击规划，不断调整攻击行为组合，直至实现攻击意图。

理性攻击敌手行为一般符合本文变迁模型，这也是攻击者所期望目标，可从著名黑客技术网站（www.phrack.org）公布论文验证。



6.网络敌手模型运行流程

本文中网络敌手模型由三个子模型组成，攻击过程中三个模型相互作用，相互影响，敌手根据新的信息更新模型中变量，模型运行流程如右图所示。

网络敌手模型实例

我们收集分析的当前网络攻击案例及常见攻击工具，根据本文所提出网络敌手模型来分别解析。

4.1网络敌手心智模型实例

下表是关于网络敌手心智状态数据实例，数据来自Michael Erbschloe的《信息战》一书所提供资料。

敌手名称	类型	愿望	意图	信念	知识	目标
Eugene \E.Kashpureff	破坏者	挫败竞争对手	干扰服务	强	中	InterNIC
Curador	黑客	炫耀	窃取信用卡号	中	高	商业网站
Kevin Mitnick	黑客	挑战	反取证	强	高	公司、大学、电信等系统
MosthateD	黑客	经济利益	电信诈骗	中	高	
Eric Burns	黑客	攻击方法指导	网页重定向	高	高	FairFax、NATO、USIA等系统
Ikenna Iffih	黑客	非法访问	非法访问	高	高	ZMOS、NASA系统

4.2网络敌手攻击决策模型实例

我们知道网络攻击的成功率随机性较大，一个复杂的攻击过程往往由若干个小的攻击环节组成，因此，攻击决策是网络敌手实施攻击的重要环节。本文实例之一就是分布式拒绝服务攻击，攻击者为了提高拒绝服务攻击的成功率，需要控制成百上千的被入侵主机。整个攻击规划可以分为以下五个步骤：第一步，通过探测扫描大量主机，寻找可以进行攻击的目标；第二步，攻击有安全漏洞的主机，并设法获取控制权；第三步，在已攻击成功的主机中安装客户端攻击程序；第四步，利用已攻击成攻的主机继续进行扫描和攻击。第五步，当攻击客户端达到一定的数目后，攻击者在主控端给客户端攻击程序发布向特定目标进行攻击的命令。

从分布式拒绝服务攻击案例来看，攻击者进行大型或复杂的攻击之前，首先需要利用已攻击成功的主机，时机成熟后再向最终的目标发起攻击。从这一点上来说，大型或复杂的攻击并不能一步到位，网络敌手需要做好攻击规划，才能实现其攻击意图。

4.3网络敌手攻击行为变迁模型实例

网络敌手攻击行为的变化取决于敌手的心智状态、攻击决策规划，下面以网络蠕虫为实例，分析敌手行为变迁。网络蠕虫将是一种智能化、自动化的攻击载体，它会主动扫描和探测网络上存在服务漏洞的节点主机，一旦渗透成功，会自我复制许多副本，通过局域网、国际互联网或者电子邮件从一个节点传播到另外一个节点。

结束语

本文主要工作在于形式化刻画了网络敌手心智、决策规划及行为特征模型，并分析各子模型相互关系。同时，提出刻画的网络敌手模型参数及方法。研究网络敌手模型的目的在于作到知己知彼，网络安全防范不仅要从正面去进行防御，更要从反面入手，掌握网络敌手模型，设计好的信息对抗系统。面对日益发展的网络系统，规模越来越大，管理复杂，松散组织方式，网络敌手容易隐藏踪迹。网络敌手的意图、策略及行为将会更加复杂多变。假若一个网络攻击者开发一个能够复制传播的攻击程序，利用其所控制的网络计算机，然后发布协同命令攻击特定的目标，很难设想攻击情形如何。我们预测，黑客攻击程序和计算机病毒传播机制的密切结合将会是新的网络攻击特点，网络敌手的能力将会增强。

（责编　杨晨）

作者简介：蒋建春，博士，单位：中国科学院软件研究所；文伟平，博士，单位：北京大学。