网络蠕虫实验环境构建技术研究

况晓辉[1,2] 黄敏桓[1,3] 许 飞[1]

(北京系统工程研究所信息系统安全技术国家级重点实验室 北京 100101)[1]

(装备指挥技术学院试验指挥系八队 北京 101400)[2] (清华大学计算机科学与技术系 北京 100084)[3]

摘 要 网络蠕虫实验环境可以为蠕虫研究提供有效的实验分析支持。在系统分析解析模型、报文级模拟、网络仿真、混合方法等蠕虫仿真环境构建技术的基础上,提出了虚实结合的蠕虫仿真模型。该模型综合了报文级模拟和网络仿真技术的优点,具有良好的扩展性和逼真度,为构建灵活可扩展的蠕虫实验环境奠定了重要基础。

关键词 网络蠕虫,实验环境,蠕虫模型

中图法分类号 TP393 **文献标识码** A

Research on Network Worm Test-bed

KUANG Xiao-hui[1,2] HUANG Min-huan[1,3] XU Fei[1]

(National Key Laboratory of Science and Technology on Information System Security, Beijing Institute of System Engineering, Beijing 100101, China)[1]

(Academy of Equipment Command & Technology, Beijing 101400, China)[2]

(Department of Computer Science & Technology, Tshinghua University, Beijing 100084, China)[3]

Abstract Worm validation experiments are very important for worm research. Based on analysis of experiment environment construct method including mathematical modeling, simulation, emulation, test bed and hybrid modeling, an new worm model named virtual and real hybrid network worm emulation model was proposed. Because of combining simulation with emulation, the model balanced the fidelity and scalability very well.

Keywords Network worm, Experiment environment, Worm model

1 前言

随着网络技术的发展,网络中存在的各种问题相继出现,在所出现的各种问题中,网络蠕虫的大面积流行对网络资源和主机造成的损失日趋扩大。2001年爆发的CodeRed蠕虫在14个小时内感染了359000台主机,造成26亿多美元的损失[1],而2003年爆发的Slammer蠕虫10分钟内就感染了超过75000台主机。随着蠕虫危害的不断扩大,网络蠕虫的机理和防御机制逐渐成为网络安全领域的研究重点。

由于蠕虫具有巨大的破坏性和不可控性,在真实网络中进行蠕虫实验存在巨大的风险。而网络蠕虫实验环境可以为蠕虫研究提供有效的实验分析支持,所以日益受到相关研究人员的关注,并逐渐成为蠕虫研究的一项重要内容[2]。网络蠕虫实验环境的应用可分为以下3方面:

1)蠕虫监测数据分析。监测数据分析,是了解相应蠕虫传播特性最为直接的一个途径,具有很高的真实性。人们在蠕虫爆发过程中可以监测到大量的蠕虫爆发实测数据,但是监测数据只是蠕虫传播过程中一个侧面的静态快照,很难从这些信息中得到蠕虫爆发过程的全部信息。利用蠕虫实验环境可以根据这些监测数据,再现蠕虫爆发时的场景,从而可以全面地获取监测数据背后的蠕虫传播信息。此外,还可以根据需要修改实验参数,充分地分析各种不同因素与蠕虫传播特性的关系,对蠕虫传播特性进行更为深入的研究。

2)蠕虫传播特性分析与评估。对于处理理论研究阶段的蠕虫,没有原始的实际测量数据。同时在真实环境中实现此类蠕虫比较困难,风险巨大,所以很难对其传播特性进行准确的分析和预测。蠕虫实验环境可以为这类蠕虫提供一个安全、有效的实验环境。研究人员可以动态地模拟其传播过程,从而对这类蠕虫的传播特性有一个科学的预测。

3)蠕虫检测防御措施的有效性分析评估。蠕虫检测防御措施的设计、分析、验证以及改进的各个环节都离不开充分的实验验证。蠕虫实验环境可以为蠕虫检测防御策略验证实验提供一个灵活、有效、安全的实验平台。通过实验,可以为检测防御策略配置合适的参数,可以在误报率、漏报率、配置有效性、响应时间、控制效果等方面,对蠕虫检测防御策略进行全面系统的评价。此外,蠕虫实验环境还可以方便地改变实验场景,以检测蠕虫防御策略中可能存在的漏洞和不足,并在分析这些不足的基础上,提出相应的改进策略。

本文在系统分析各类蠕虫仿真环境构建技术的基础上,针对已有实验环境在扩展性和逼真度平衡、评估蠕虫防御机

到稿日期:2009-08-26 返修日期:2009-10-09 本文受国家863计划(2009AA01Z421)资助。

况晓辉(1975—),男,副研究员,主要研究方向为计算机网络技术与信息安全等,E-mail:xiaohui_kuang@163.com;黄敏桓(1971—),男,研究员,主要研究方向为网络技术与信息安全等;许 飞(1981—),男,助理研究员,主要研究方向为计算机网络技术。

制有效性等方面的不足，将报文级模拟和网络仿真等技术有机结合，提出了虚实结合的蠕虫仿真模型。该模型具有较高的逼真度和良好的扩展性，且可直接部署蠕虫网络防御机制，为构建支持大规模网络蠕虫的传播机理研究和防御机制评估的实验环境奠定了重要基础。

2 网络蠕虫实验环境构建技术

网络蠕虫实验环境构建技术可分为实验床、网络仿真、报文级模拟、解析模型以及混合仿真等类型，基于不同构建技术建立的实验环境，在逼真度和扩展性上存在很大差异，如图1所示。

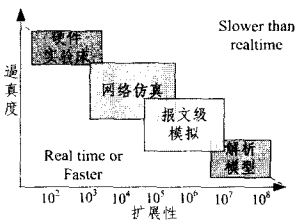

图1 实验环境构建技术

・解析模型

解析模型是研究蠕虫传播特性最常用的一种方法。它基于一组微分方程或者离散的递归表达式构建蠕虫实验环境，描述蠕虫在网络中传播的动态过程。目前主要的解析模型方法有连续数学模型和离散数学模型两类。前者是采用微分方程来模拟蠕虫传播的过程，如SI，SIR，KM以及双因素模型[3]等；后者将蠕虫传播过程分解为若干个时间片，蠕虫可以在单个时间片内完成节点感染动作，从而将蠕虫传播过程离散化，如AAWP模型[4]。

解析模型的优势在于可扩展性强，能够模拟几百万至几千万个节点的网络。然而，该模型仅集中于蠕虫自身的传播机理，忽略了网络中的节点数、链路状态、业务流量的产生和变化以及拓扑结构等特性，因此逼真度不足，难以支持对蠕虫破坏性和防御机制的分析评估。

・报文级模拟

报文级模拟技术建立在网络模拟技术基础上，它将蠕虫数据包传播看作一种特殊的网络流量，通过模拟蠕虫数据包在网络环境中的转发过程，模拟蠕虫在特定网络环境下的传播过程。基于报文级模拟技术构建的实验环境可以准确模拟蠕虫数据包转发发生的随机性，体现网络拓扑结构的多样性。目前，在许多著名的网络模拟器上均开展了蠕虫仿真技术研究工作。Li Lunquan等人在NS2基础上实现了DN-AN混合数据包仿真模型，支持对slammer蠕虫的模拟[5]。George Riley等人在GTNets网络仿真器的基础上实现了一个完整的数据包级蠕虫仿真系统[6]。邓洁等在SSFNet的基础上设计并实现了蠕虫传播细粒度建模与仿真[7]。

基于报文级模拟技术构建的蠕虫实验环境具有良好的扩展性，单个模拟实例可模拟成百上千个节点构成的网络。采用分布式模拟技术，由多个模拟实例相互协作，还可构建规模更大的实验环境，如PDNS[8]和GTNetS[9]。同时，虽然抽象级别不同，但是基于报文级模拟技术构建的实验环境通常均考虑了带宽、拥塞、延迟、拓扑、流量和丢包率等网络特性，在逼真度上优于解析模型。但是，报文级模拟对端系统的逼真度支持不足，且蠕虫防御机制和工具需要在模拟器上重新开发才能进行评估，因此难以有效支持针对端系统的复杂交互式蠕虫破坏性的分析评估，不能直接对蠕虫防御工具和软件的有效性进行分析评估。

・网络仿真

基于网络仿真技术构建的实验环境在真实主机和实验节点间建立一一对应的关系，以获得精确的结果。网络仿真还可以细分为实验床或层叠网两种类型，前者如Emulab网络实验床[10]和DETER[11]等；而Planetlab[12]则是典型的基于层叠网技术构建的实验环境。上述两类实验环境均可运行真实的蠕虫代码，观测蠕虫流量，分析蠕虫的传播过程。Matthew M. Williamson等人基于DETER提供的实验环境构建了一个小型实验网络，分析了Virus-Throttle蠕虫防御策略对Nimda蠕虫代码的防御效果[13]。

基于网络仿真技术构建的实验环境在网络特性和端系统等层次均具有较高的逼真度，且能够直接评估蠕虫防御工具和机制针对真实蠕虫的有效性。但是，此类实验环境的规模受限于实验床的节点数量，其仿真规模远远不能满足实际的蠕虫传播研究需要。同时，系统规模与蠕虫传播特性之间未存在严格的线性关系，很难从小规模的仿真结果直接推断出大规模蠕虫传播特性[2]。因此，仅采用网络仿真方法来构建的实验环境无法满足蠕虫仿真的规模需求。

・硬件实验床

硬件实验床是真实的网络系统，即采用与真实网络一致的软硬件构建的实验环境。由于网络蠕虫对于实验环境规模需求较大，无法基于硬件实验床方法构建，因此目前没有针对网络蠕虫的硬件实验床。

・混合方法

为平衡扩展性和逼真度间的矛盾，人们提出了多种混合方法来构建蠕虫实验环境。从蠕虫实验环境构建技术可看出，可行的混合方法可分为报文级模拟和解析模型相结合的混合方法以及网络仿真和报文级模拟相结合的混合方法。对于前者，M. Liljenstam等人在SSFNET仿真器的基础上实现了基于报文级模拟与解析模型相结合的混合蠕虫模拟实验环境[14]。王跃武等在M. Liljenstam等人工作的基础上，提出了基于选择抽象的蠕虫仿真模型，以实现仿真保真度和仿真规模间的灵活调节，从而提高仿真资源的有效利用[2]。

采用报文级模拟和解析模型相结合的混合方法构建的蠕虫实验环境具有良好的扩展性和模拟效率，但是未能解决端系统逼真度以及蠕虫防御机制分析评估问题。网络仿真和报文级模拟相结合的混合方法既能够发挥报文级模拟技术扩展性强的优点，又可利用网络仿真技术逼真度高的优势，且构建的实验环境采用硬件在回路的仿真思想，可与真实的网络设备和软件交互，因此可直接分析评估蠕虫防御机制的有效性。Deterlab以及emulab等均采用此思想构建了大规模网络实验环境，但是目前没有基于网络仿真和报文级模拟相结合的蠕虫仿真实验环境。

2.1 对比分析

根据对不同实验环境构建技术的分析，可发现报文级模拟和网络仿真相结合的混合方法在网络特性、节点特性、蠕虫特性和环境特性等方面具有明显优势。将各种构建技术对特性的支持分为4个档次(√表示能够充分支持该特性，△表示

可部分支持该特性，？表示难以支持该特性，×表示无法支持该特性），各种实验环境构建技术的对比如表 1 所列。

表 1　仿真方法的比较

特性	方法	解析模型	报文级模拟	网络仿真	解析模型 & 报文级模拟	报文级模拟 & 网络仿真
网络特性	网络带宽	？	√	√	√	√
	网络拥塞	×	√	√	√	√
	网络延迟	△	√	√	√	√
	网络拓扑	？	√	√	√	√
	网络流量	？	√	√	√	√
	网络协议	？	√	√	√	√
	背景流量	×	√	√	√	√
主机特性	数据包丢失	×	√	√	√	√
	系统补丁机制	×	×	√	×	√
	安全防护工具	×	×	√	×	√
	蠕虫清除机制	？	×	√	？	√
蠕虫特性	蠕虫大小	√	√	√	√	√
	传输协议	？	√	√	√	√
	扫描速率	√	△	√	√	√
	感染时间	√	√	√	√	√
	感染延迟	√	√	√	√	√
	传输端口	？	√	√	√	√
	传输方式	√	△	√	√	√
	扫描策略	√	△	√	√	√
环境特性	扩展性	√	√	×	√	√
	逼真度	×	△	√	△	√
	防御机制评估	×	△	△	△	√

3　虚实结合的网络蠕虫仿真模型

虚实结合的网络蠕虫仿真模型采用网络仿真和报文级模拟相结合的方法，包含基于试验床的网络仿真、报文级模拟和蠕虫行为模型 3 个层次，如图 2 所示。

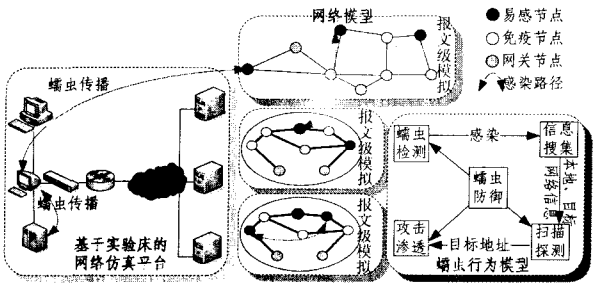

图 2　虚实结合的网络蠕虫仿真模型

基于实验床的网络仿真平台由各类网络设备、服务器构成。服务器可根据实验需要配置各类服务器和端系统，如具有特定脆弱性的主机，可以支持真实网络蠕虫传播；也可加载蠕虫模拟程序或软路由程序成为实验节点；报文级模拟技术用于在单个实验节点上构建支持蠕虫传播的虚拟网络，为网络蠕虫仿真模型提供可扩展性。虚拟网络由节点和链路构成，其中节点包括网关节点、主机节点两种。网关节点负责报文的转发，包括虚拟网络报文的转发以及虚拟网络与真实环境间报文的转换和路由；主机节点根据易感节点分布策略是否加载蠕虫仿真模型，分为易感节点和免疫节点。链路由丢包率、延迟、带宽等链路特性和队列模型构成，可对不同类型的链路进行仿真。

蠕虫行为模型在虚拟网络的主机节点上仿真蠕虫的行为。根据网络蠕虫功能结构分析，蠕虫行为模型包含蠕虫监测模块、信息搜集模块、扫描探测模块、攻击渗透模块以及蠕虫防御模块 5 部分。蠕虫监测模块监视到达目标节点的报文，分析其是否包含蠕虫的特征码，判断节点是否遭受蠕虫攻击。信息搜集模块实现对本地或者目标网络信息的搜集，内容包括节点信息、用户信息、邮件列表、网络拓扑结构等；扫描探测模块基于信息搜集模块获得的数据以及扫描策略确定渗透攻击的目标。攻击渗透模块利用扫描探测模块确定的目标加载攻击载荷，实施蠕虫渗透攻击；蠕虫防御模块反映外界因素和蠕虫对抗措施对蠕虫的影响，例如各 ISP 节点或用户的防御措施等。

通过网关节点的虚拟报文实时转换以及蠕虫仿真模型的监测机制，虚实结合的网络蠕虫仿真模型可实现虚拟网络和仿真平台的蠕虫传播，即蠕虫不仅可在实验床中真实的设备间传播，还可在真实网络和虚拟网络的易感节点间传播以及在虚拟网络内部易感节点间传播。

由于有机融合了网络仿真和报文级模拟方法，虚实结合的网络蠕虫仿真模型较好地平衡了逼真度和扩展性两方面的矛盾，实验人员可根据实验环境的硬件资源和研究需要，灵活地确定二者间的比重。此外，该模型由于支持虚实结合的蠕虫仿真，因此可分析真实蠕虫的传播特性和对实验网络的影响，并可直接部署蠕虫网络防御机制，评估其效能。同时，该模型还支持不同网络模拟实例间的蠕虫传播，可通过分布式模拟的方法分析自定义蠕虫的传播特性及其对于大规模网络的影响。

结束语　本文在全面分析目前蠕虫实验环境构建技术相关研究的基础上，提出了一个虚实结合的网络蠕虫仿真模型。该模型有机融合了网络仿真和报文级模拟方法的优点，在逼真度和扩展性的平衡方面优于已有的网络蠕虫实验环境。下一步将研究仿真模型的实现技术，构建虚实结合的网络蠕虫实验环境，研究蠕虫传播机理和防御机制的分析评估技术。

参 考 文 献

[1] The Spread of the Code-Red Worm (CRv2)[EB/OL]. http://www.caida.org/research/security/code-red/coderedv2_analysis.xml

[2] 王跃武. 网络蠕虫仿真技术研究[D]. 北京：中科院软件所，2008

[3] Zou C C, Gong W, Towsley D. Code Red worm propagation modeling and analysis[C]//Proc. of the 9th ACM Symp. on Computer and Communication Security. Washington, 2002: 138-147

[4] Nicol D M. The Impact of Stochastic Variance on Worm Propagation and Detection[C]//Proc. of the 2006 ACM Workshop on Rapid Malcode, Alexandria Virginia. NY: Publications Dept., ACM, Inc, 2003: 11-18

[5] Li Lunquan, Jiwasurat S, Liu Peng, George Kesidis § Emulation of "single-packet" UDP Scanning Worms in Large Enterprises

[6] Riley G F, Sharif M I, Lee W. Simulating Internet Worms[C]// 12th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS). Oct. 2004

[7] 邓洁，段海新. 基于 SSFNet 的蠕虫传播细粒度建模与仿真[J]. 小型微型计算机系统，2008，29(1)

[8] Kalyan S. Perumalla Srikanth, Sundaragopalan. High-Fidelity Modeling of Computer Network Worms[C]// 20th Annual Computer Security Applications Conference (ACSAC'04). 2004: 126-135

（下转第 73 页）

$$B_3=\begin{bmatrix}1 & 2 & 5\\ 1/2 & 1 & 2\\ 1/5 & 1/2 & 1\end{bmatrix}$$

然后计算第二层次的相对权重。首先求出特征向量 $M_1=(0.246,0.710,2.047)^T$，$M_2=(0.544,0.215,2.293)^T$，$M_3=(1.791,0.831,0.386)^T$。对 M_1,M_2,M_3 进行归一化，得 $W_1=(0.082,0.236,0.682)^T$，$W_2=(0.180,0.072,0.748)^T$，$W_3=(0.595,0.277,0.128)^T$。再计算 $\lambda_{max1}=3.002$，$\lambda_{max2}=3.029$，$\lambda_{max3}=3.006$。一致性指标 $C.I._1=0.001<0.1$，$C.I._2=0.015<0.1$，$C.I._3=0.015<0.1$。表明上述判断矩阵一致性均能成立。W_1,W_2,W_3 中的元素即为各指标相对于上一层的权值。

攻击效果(y_1,y_2,y_3)以上一层 Z 为依据的判断矩阵是

$$A=\begin{bmatrix}1 & 2 & 6\\ 1/2 & 1 & 4\\ 1/6 & 1/4 & 1\end{bmatrix}$$

由 AHP 法计算得特征向量 $M=(1.769,0.974,0.268)^T$，对 M 进行归一化，得 $W=(0.587,0.324,0.089)^T$。再计算 $\lambda_{max}=3.029$，一致性指标 $C.I.=0.0045<0.1$，表明判断矩阵一致性成立。W 中的元素即为各攻击效果的权值。

4.3 熵差计算与分析

进行不同频率的路由攻击时，对系统受攻击后的指标进行测量，将结果代入式(6)，进行熵差计算。计算结果如表 3 所列。其中 f 表示攻击频率(次/100s)，$A=\{a_1,a_2,a_3\}$ 表示攻击的类型，ΔS 表示受攻击后与受攻击前的路由安全熵差，由各层指标加权计算而得。

表 3　受到不同频率路由攻击时的熵差变化

ΔS \ f / a	10	20	30	40	50	60	70	80	90	100
a_1	0.0740	0.2345	0.4150	0.6439	0.8625	1.0291	1.3219	1.7370	2.1520	3.0589
a_2	0.0291	0.2176	0.3959	0.6215	0.8890	1.0000	1.2863	1.6891	2.1203	2.7370
a_3	0.1203	0.2863	0.4639	0.7131	0.9434	1.1520	1.3959	1.8110	2.4739	3.3219

通过计算，表 3 给出了受到不同频率的路由攻击时的安全熵差的变化。从表中可以看出，遭受每一种攻击时，熵差 ΔS 都随着攻击频率的增加而增加。由 2.2 节可知，熵差越大，说明网络受到路由攻击之后对路由造成的安全危害越大。因此，计算结果表明，随着攻击频率的增加，系统的安全性降低。使用熵差 ΔS 可以客观真实地反映 WSN 的路由协议在遭受攻击时的安全性。根据计算结果，可以将路由的安全性分为以下 3 个等级：当 $0\leqslant\Delta S<0.5$ 时，系统为较安全系统；当 $0.5\leqslant\Delta S<1.8$ 时，系统为一般安全系统；当 $\Delta S\geqslant1.8$ 时，系统为不安全系统。

本文针对无线传感器网络路由协议的具体情形，首先以路由的安全性为总体依据，建立安全度量层次，并采用层次分析法确立评估的具体指标。通过比较各个指标的重要性，确定其权值。在此基础上针对不同路由协议给出不同的主要评估指标，从而减少了指标确定的主观因素。其次，利用 Monte Carlo 概率统计的方法对安全度进行归一化，通过多次测量进一步提高了结果的准确性，而且可以避免重复的漏洞检测工作，节省了大量时间和空间，并能够通过概率变化预测系统的安全态势。利用“网络安全熵”的概念计算路由的整体安全水平，计算结果对安全性进行了精确的定量描述，并区分了路由安全等级。利用“熵”的概念，进一步减少了其他因素的干扰，且避免了大量的样本空间。

结束语　本文提出并验证了基于攻击效果的 WSN 路由安全评估模型。结合统计概率和相关权值分析，提出了“网络安全熵”的概念，从而使得模型能够对无线传感器网络的安全性进行客观真实的量化评价，有助于提高无线传感器网络的安全性，从而提高系统应对各种攻击的能力，并能对发现、反击敌方的恶意攻击起到指导作用。实验结果表明，本模型安全评价合理，有较高的应用价值。

参 考 文 献

[1] 裴庆祺，沈玉龙，马建峰. 无线传感器网络安全技术综述[J]. 通信学报，2007，28(8)：113-122

[2] Zhang Y R，Xian M，Wang G Y. A quantitative evaluation technique of attack effect of computer network based on network entropy[J]. Journal on Communications，2004，25(11)：158-165

[3] 肖道举，杨素娟，周开锋，等. 网络安全评估模型研究[J]. 华中科技大学学报，2002，30(4)：37-39

[4] Krontiris I，Dimitriou T，Freiling F C. Towards intrusion detection in wireless sensor networks[C]// Proceedings of the 13th European Wireless Conference. Paris，France，2007

[5] 赵冬梅，马建峰，王跃生. 信息系统的模糊风险评估模型[J]. 通信学报，2007，28(4)：51-56，64

[6] 张小川，李祖枢. 基于人工生命行为选择的智能体决策的研究[J]. 计算机科学，2007，34(5)：213-214，251

[7] Maarouf I. Efficient monitoring approach for reputation system-based trust-aware routing in wireless sensor networks[J]. IET Communications，2009，3(5)：846-858

[8] Jaquith A. Security Metrics：Replacing Fear，Uncertainty，and Doubt[M]. 李冬冬，韦荣，译. 北京：电子工业出版社，2007：22

[9] 包秀国，胡铭曾，张宏莉，等. 两种网络安全管理系统的生存性定量分析方法[J]. 通信学报，2004，25(9)：34-41

(上接第 56 页)

[9] Sewani A. Packet Level Worm Simulation and Analysis[OL]. http://www.cs.berkeley.edu/~anil/papers/worm_sim.ps

[10] http://www.emulab.net

[11] http://www.isi.edu/deter.

[12] http://www.planet-lab.org

[13] Twycross J，Williamson M M. Implementing and testing a Virus Throttle[C]// Proceedings of 12th USENIX Security Symposium. August 2003

[14] Liljenstam M，Yuan Y，Premore B，et al. A mixed abstraction level simulation model of large-scale internet worm infestations [C]//MASCOTS. IEEE，October 2002