doi:10.11959/j.issn.1000-436x.2017123

网络协议隐形攻击行为的聚类感知挖掘

胡燕京[1,2]，裴庆祺[2]

（1. 武警工程大学网络与信息安全武警部队重点实验室，陕西 西安 710086；
2. 西安电子科技大学综合业务网理论及关键技术国家重点实验室，陕西 西安 710071）

摘 要：深藏在网络协议中的隐形攻击行为日益成为网络安全面临的新挑战。针对现有协议逆向分析方法在协议行为分析特别是隐形攻击行为挖掘方面的不足，提出了一种新颖的指令聚类感知挖掘方法。通过抽取协议的行为指令序列，利用指令聚类算法对所有的行为指令序列进行聚类分析，根据行为距离的计算结果，从大量未知协议程序中快速准确地挖掘出隐形攻击行为指令序列。将动态污点分析和指令聚类分析相结合，在自主研发的虚拟分析平台 HiddenDisc 上分析了 1 297 个协议样本，成功挖掘出 193 个隐形攻击行为，自动分析和手动分析的结果完全一致。实验结果表明，该方案在效率和准确性方面对协议隐形攻击行为的感知挖掘都是理想的。

关键词：协议逆向分析；隐形攻击行为；指令聚类

中图分类号：TP393 **文献标识码**：A

Clustering perception mining of network protocol's stealth attack behavior

HU Yan-jing[1,2], PEI Qing-qi[2]

(1. Network and Information Security Key Laboratory, Engineering University of CAPF, Xi'an 710086, China;
2. State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China)

Abstract: Deep stealth attack behavior in the network protocol becomes a new challenge to network security. In view of the shortcomings of the existing protocol reverse methods in the analysis of protocol behavior, especially the stealth attack behavior mining, a novel instruction clustering perception mining algorithm was proposed. By extracting the protocol's behavior instruction sequences，and clustering analysis of all the behavior instruction sequences using the instruction clustering algorithm, the stealth attack behavior instruction sequences can be mined quickly and accurately from a large number of unknown protocol programs according to the calculation results of the behavior distance. Combining dynamic taint analysis with instruction clustering analysis,1 297 protocol samples were analyzed in the virtual analysis platform hidden disc which was developed independently，and 193 stealth attack behaviors were successfully mined，the results of automatic analysis and manual analysis were completely consistent. Experimental results show that，the solution is ideal for perception mining the protocol's stealth attack behavior in terms of efficiency and accuracy.

Key words: protocol reverse analysis，stealth attack behavior，instruction clustering

1 引言

网络协议的隐形攻击行为是指通过网络协议成功实现攻击目标而不被现有安防设备和安防技术所感知的攻击行为。隐形攻击行为通常很难被动态分析捕获到，只有在特定条件下才会执行。由于采用自我保护手段，静态分析对它同样难以识别。该行为具有潜伏期长、执行期短、恶意性不显著但潜在危害大等特点。近年来，网络协议的隐形行为技术得到了长足的发展，这就为网络隐形攻击的实施开辟了新的道路。网络协议的所有行为均不会超出协议规定的范畴，隐形攻击行为同样也不例外，

收稿日期：2016-11-02；**修回日期**：2017-04-18
基金项目：国家自然科学基金资助项目（No.61373170, No.61402530, No.61309022, No.61309008）
Foundation Item: The National Natural Science Foundation of China (No.61373170, No.61402530, No.61309022, No.61309008)

因此，从实现协议的二进制程序代码中逆向感知挖掘出隐形攻击行为成为一个根本的技术路线。目前，发现的隐形攻击行为大体分为2种模式：一是通过精心设计的隐匿功能秘密窃取目标网络的高价值信息；二是利用已掌握的协议漏洞，将预先设计好的隐匿功能静默嵌入到正常协议之中，长期潜伏，只在特殊条件具备时才触发执行，一旦完成任务又继续潜伏，从而长久、巧妙地控制目标主机而不被发现。然而感知和挖掘网络协议的隐形攻击行为却面临极大的困难，因为隐形行为种类繁多，性质和目的各异，没有一个统一固定的行为模式可以作为检测的依据，而日益增长的隐蔽性、顽健性和可生存性，使传统的分析、追踪和识别变得更加困难。

网络协议的行为安全是网络空间安全的基础和保障，然而，目前，对协议行为并没有一个完整清晰的认识，协议行为的定义、挖掘和分析方法亟待探讨，至今仍无一套协议行为模型或协议行为分析方法能够保证网络协议的安全运行。因此，本文从协议逆向分析出发，探索和尝试协议行为的分析方法，研究正常行为和隐形行为的基本特征，以协议行为的分析挖掘为突破口，对未知协议的隐形攻击行为分析方法展开研究。本文运用动态污点分析[1]、静态分析[2]、指令聚类分析和关联分析等技术手段，对未知网络协议的隐形攻击行为进行逆向分析[3]和感知挖掘，在准确性、效率和可靠性方面均取得了较好的结果。

2 相关工作

传统的协议逆向分析只关注未知协议的语法、消息格式的推断和恢复等问题[4~7]，并不涉及协议行为的研究，而协议的行为却和网络安全息息相关。在日益严峻的网络安全形势下，协议行为的逆向分析被提出，主要以协议程序的功能性指令序列为数据源，从虚拟仿真平台上监控协议的指令序列入手，分析表示协议行为的指令序列，对协议的行为进行评估，着力应对由未知协议的异常行为造成的隐性安全威胁。协议软件是协议最终的实现，对协议行为的分析可转化为对协议软件的分析。协议行为分析是协议逆向分析的重要组成部分，在网络安全中发挥着基础性的作用，虽然这一概念很少明确提出，但行为分析技术在入侵检测、入侵防御、恶意代码分析和安全审计等工作中已经得到应用，成为传统网络安全设备的有益补充。目前，与协议行为分析密切相关的概念并没有完全统一，主要研究包括协议逆向工程、恶意软件行为分析、网络安全审计和网络行为分析等技术。

协议行为的表示粒度各有不同，前人多用系统调用序列来表示协议软件的行为[8]。Forrest 等在1996年提出了基于系统调用的入侵检测模型，成为这一时期行为分析的典型工作。由于系统调用的基础性作用，进程系统调用的相关属性成为用来描述进程行为，实施入侵检测的重要依据。苏璞睿等[9]提出了一种基于可执行文件静态分析的进程行为建模方法，通过静态分析应用程序可执行代码，建立进程运行过程中可能的系统调用序列集合,以该集合为基础，对进程实施监控，进程执行过程中，出现任何不在该集合中的系统调用片断均认为行为异常，即发生了攻击。该模型主要有3个特点：1) 分析过程不依赖于任何源程序，可对各类应用程序实施监控；2) 不会出现任何误报；3) 具有较强的抵抗“模仿攻击”的能力。受Forrest等工作的启发，后续相继出现了大量优秀成果，包括Var-gram模型、FSA 模型、Vt-Path 模型等。但这些模型都太依赖于系统调用，假如攻击者用自己的代码代替系统调用的功能，则这些方案在行为建模、检测准确性、检测能力和实用性方面就需要进行重大改进和提高才能应对这一艰巨的挑战。

目前，学术界对协议的行为并没有一个完整、清晰的认识，现有的网络协议理论及工具还不能保证协议的异常行为或隐形攻击行为产生时能积极应对和有效管控，对协议隐形行为的检测机制尚待探讨，至今仍无一套表示和分析协议行为的模型或方法。协议行为特别是隐形攻击行为分析不同于传统的协议逆向，它关注的是协议的行为信息，需要收集和分析协议运行时的各种数据，既需要协议消息流，也需要指令数据流，为了有效挖掘协议的隐形攻击行为，既要动态分析也要静态分析，还需要聚类挖掘和关联分析等领域的技术，因此，协议隐形攻击行为的逆向分析需要投入更多的精力进行专门研究。

3 协议隐形攻击行为的聚类感知挖掘方案

3.1 问题描述和定义

协议设计者的目标最终都是通过协议的行为表现出来的，协议设计者的意图和协议行为之间具

有可信的对应关系。协议的正常行为是各种网络应用的正常表现，而协议的隐形攻击行为则是协议设计者特殊目标的表现形式。这些特殊目标通常不愿意被公开表现出来，往往采用隐匿、隐蔽、隐藏等方式加以保护，以逃避安全分析人员的分析和追踪，采用自迷惑技术，在不改变原始语义的前提下通过代码变换来抵御逆向分析。所以，协议的隐形攻击行为都是通过精心的设计、巧妙的躲藏来完成使命的。从本文捕获到的协议样本中可以发现，这些有意设计的隐形攻击行为主要有秘密扫描硬盘文件、非法下载、植入特殊代码、隐私侦听和拷贝用户信息等，具有现实或潜在的恶意性。

为了对未知协议二进制程序处理消息的过程细节进行研究，本文开发了一个虚拟分析平台 HiddenDisc (hidden behavior discovery)。该分析平台根据所有待分析协议的指令级特性进行指令聚类分析，自动将所有协议样本的行为进行分组并生成不同的行为聚类，将动态二进制分析和静态指令聚类分析相结合。为了应对未知协议潜在的隐形攻击行为，本文从一个全新的视角重新描述了协议、协议行为以及隐形攻击行为分析等问题。

网络协议的新描述。网络协议 P 可以看作是由协议程序 C 和协议消息 M 组成的集合。协议 $P=(C,M)$ 可被看作一个二元组，C 是实现协议的程序代码集合，M 是协议实体间传输的消息集合，且 M 由 C 生成。协议程序 $C=\{c_1,c_2,\cdots,c_n \mid n\in N\}$ 可被抽象为一个面向功能的指令序列 c_j 的集合，$c_0=\{i_1,i_2,\cdots,i_k\}(k>0,1\leqslant j\leqslant n)$ 是一个可执行的有序指令序列。协议消息 $M=(M_s, M_r)$也是一个二元组，M_s是发送消息的集合，而 M_r是接收消息的集合。

协议行为。协议行为 $B:M\to C$ 可被看作是从协议消息 M 到协议程序 C 的映射，一个协议的主要行为来源于消息解析过程。设 M 是协议消息的集合，则协议行为 B 可以描述为 $B_r=\text{Receive}(M)$，$B_p=\text{Process}(M)$，$B_t=\text{Trigger Behavior}(M)=B_{pub}\cup B_{dorm}$，$B_g=\text{Generate}(M)$，$B_s=\text{Send}(M)$。在行为 B_t 中，B_{pub} 是公开行为，而 B_{dorm} 是隐形行为，$B_{dorm}=\Phi$ 意味着协议没有隐形行为。各部分行为之间的关系是 $B_r\cup B_p\cup B_t\cup B_g\cup B_s\subseteq B$ 而 $B_r\cap B_p\cap B_t\cap B_g\cap B_s=\Phi$。

协议隐形攻击行为分析指在没有协议规格说明的条件下，仅获得协议程序 C 和捕获到的协议消息 M，监视和分析 C 解析 M 的过程。1) 构建协议行为指令序列；2) 根据公开行为指令序列推断潜在的触发条件；3) 通过静态聚类分析挖掘协议的隐形攻击行为，并且触发其执行作为验证。

3.2　系统框架设计

本文提出并自主研发的协议隐形攻击行为感知挖掘系统主要由 6 大功能模块组成。系统框架如图 1 所示。

1) 数据分组分析

本文把“RADCOM”硬件协议分析仪和“Wire shark”协议分析软件同时部署使用，“Wire shark”主要监视分析所有进入实验客户机的网络数据，而“RADCOM”硬件协议分析仪则为分析服务器量身打造。2 种协议分析仪同时使用，也是为了便于对比分析。数据分组分析阶段对网络原始数据进行初步识别和字段分析，同时将这些数据分组存储起来，供下一步动态污点分析使用。

2) 协议消息的动态污点分析

实例分析发现，协议的主要行为都和消息解析过程密切相关，解析消息的指令序列里蕴含着大量的协议行为信息。因此，本文把协议消息的每一个

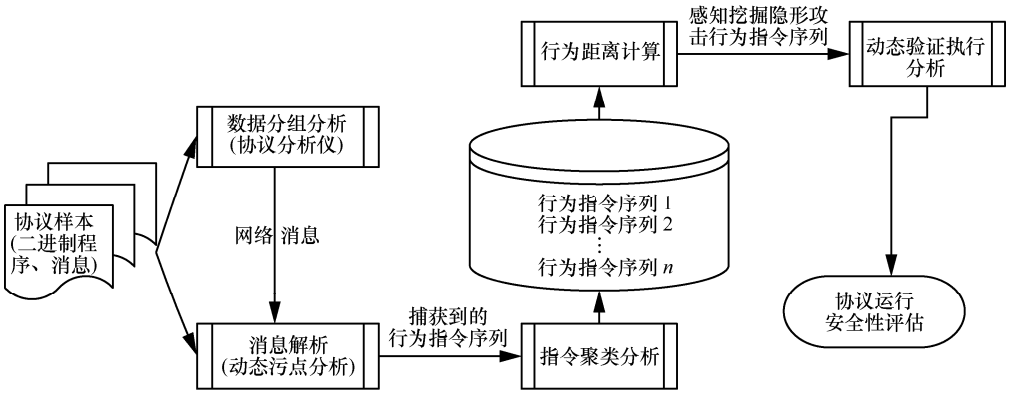


图 1　协议隐形攻击行为感知挖掘

字节都作为污点源进行标记，监控其在内存和寄存器中的传播情况，监视和记录协议程序解析污点数据的全部指令，从而捕获到机器指令级的行为指令序列。

3) 指令聚类分析

大量研究发现，相似的协议行为在指令的分布上也大体相似，如相似的指令类型、数量和调用频率等。本文将 *n*-gram 分析引入指令分析，设计出了指令聚类分析算法。根据该算法，各种协议的相似行为根据指令的分布形成同一个行为聚类，而隐形攻击行为由于在指令分布上和正常行为差距较大，从而被挖掘出来，生成不同的行为聚类。经过指令聚类分析，就可以得到一个个行为指令序列，而潜在的隐形攻击行为就包含其中。

4) 行为距离计算

将捕获到的每一个行为指令序列和已知行为进行计算，若已知行为是确定的正常行为，和其行为距离接近的行为都可视为正常行为，而行为距离偏离最大的则是被感知挖掘出的隐形攻击行为。

5) 动态验证执行分析

挖掘出的隐形攻击行为具体功能是什么，意图目的何在，还需要触发其指令序列的执行和分析来进行验证，从而掌握隐形攻击行为的特征。

6) 协议运行安全性评估

根据验证执行分析的结果，生成协议运行安全性评估报告，作为一次分析的结果和协议行为安全研究的第一手资料。

3.3 指令聚类分析方法

3.3.1 行为指令序列的抽取

近年来，本文捕获了大量各种各样的协议样本，但其中绝大多数协议的行为未知，由于多数协议都采用加密、混淆等技术手段进行反逆向分析，传统的静态分析和动态分析对隐形攻击行为的挖掘均很难奏效。本文自主研发的虚拟分析平台 HiddenDisc，将所有协议样本的全部行为指令序列抽取出来进行分析研究。HiddenDisc 是专门为自动逆向分析协议的未知行为而设计的，它以虚拟平台 TEMU 为基础进行二次开发，扩展增加了 4 大主要功能部件，如图 2 所示。

对协议样本的分析可以通过 2 条技术路线来实施，一是对二进制协议程序进行反汇编，并进一步划分成基本块；二是对协议消息的每一个字节做标记，进行污点传播分析。如果协议样本进行了加密或混淆保护，则反汇编静态分析就难以奏效了，此时就应该在动态分析环境中将协议程序执行和分析。在动态分析阶段，动态执行监视器监视协议程序解析协议消息的每一条二进制指令，同时监视污点数据（协议消息的每个字节）是如何传播的。在动态执行监视器的帮助下，和消息解析行为相关的基本块自动生成。这些原始基本块的数据可能会非常庞大，因为一些系统调用和无意义的代码块也包含其中。所以本文需要从这些庞大的原始基本块中抽取核心指令序列，以提高安全分析的效率。使用本文自主开发的用户自定义 API，就可以从原始基本块的解析过程中抽取出捕获到的行为指令序列。图 3 为从原始基本块中抽取有效行为指令序列的实例。去除指令地址、操作数、无关参数和无关指令，最终得到精简优化的只含有操作码的行为指令序列。

3.3.2 聚类感知挖掘算法

本文提出并实现了指令聚类算法，用来感知、识别和挖掘深藏在未知协议中的隐形攻击行为。指令聚类将所有的行为指令序列进行划分，每一个分组由一个或若干个行为指令序列组成，它们在指令的数量和频率上具有一定的相似性，称为一个行为聚类。由于指令是协议行为的微观表现，所以不同

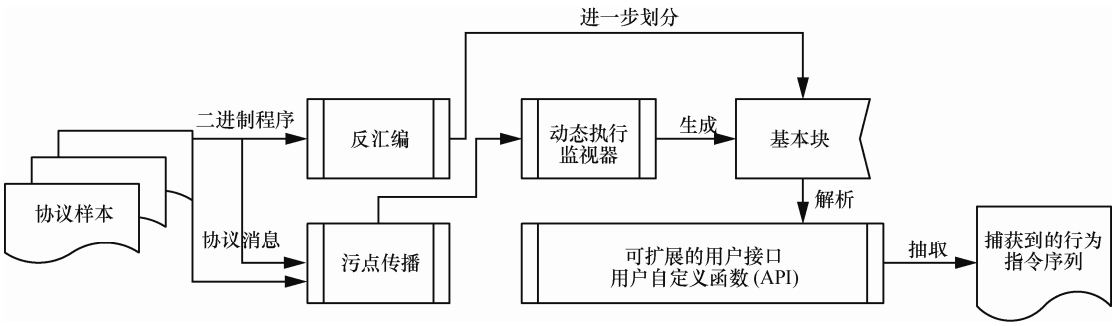


图 2 在 HiddenDisc 虚拟分析平台上捕获行为指令序列

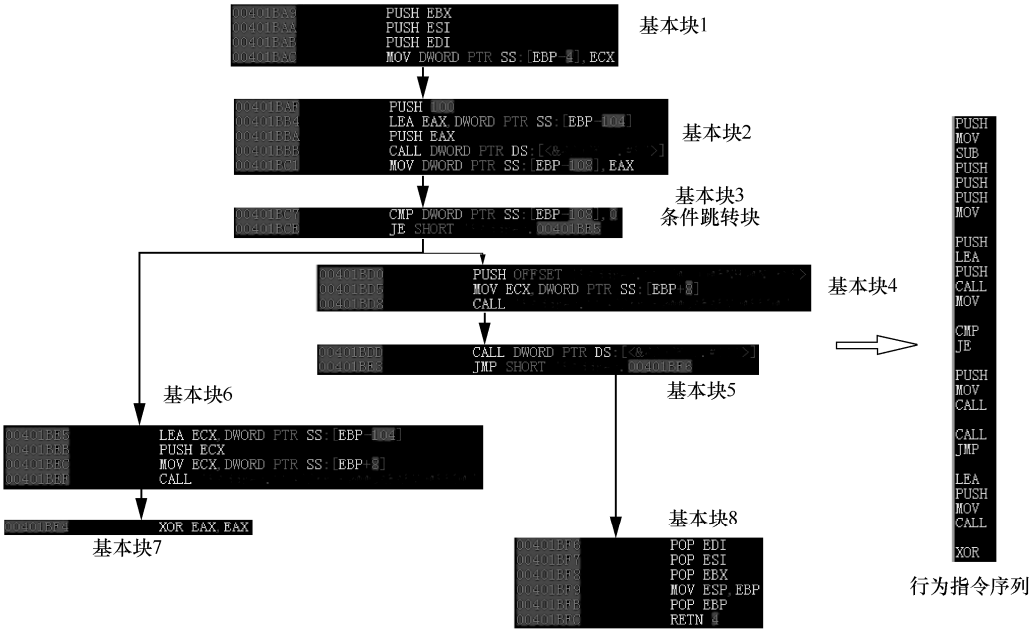


图 3　行为指令序列的生成示例

的指令序列构成不同的行为聚类。指令聚类分析模块通过大量协议样本的训练，已经积累了相当数量的正常和异常行为模式，当然还可以不断预测和挖掘新的协议行为模式。发送和接收消息是协议典型的行为模式，而且行为模式的数量也很有限，那些和正常行为偏离较大的隐形攻击行为可以通过指令聚类分析识别和挖掘出来。隐形攻击行为的聚类感知挖掘由算法 1 所示。

算法 1　协议隐形攻击行为的聚类感知挖掘

getHiddenBehavior(A,Msg)

/* 通过指令聚类分析来挖掘未知协议的隐形攻击行为*/

输入　A= $\{a_1,a_2,\cdots,a_n|\ n\in N\}$// 已经捕获到的行为指令序列

Msg //协议分析仪捕获到的消息实例

输出　隐形攻击行为集 D_S

Begin

/*指令聚类阶段*/

循环开始

CreateBehaviorCluster(A); //从协议二进制训练集中选择一个捕获到的已知行为指令序列 A 作为初始行为模式，这个被选择的行为模式生成了一个基本的行为聚类

MarkInstructions(F, C, D); //所有指令被标记为 3 种类型的基因指令：函数调用相关指令 F、条件跳转相关指令 C 以及数据处理相关指令 D

N-gramAnalyzeInstructions(F,C,D);// 基于基因指令分布的相似性，对用基因指令标记后的所有指令（用 F、C 和 D 进行标记）进行 n-gram 分析，生成不同的行为聚类

Instruction Clustering();

CreateBehaviorClusters();// 将所有挖掘到的新的指令序列生成新的行为聚类

反复迭代，直到没有行为指令序列被挖掘出为止；

循环结束

将所有指令序列 $C=\{c_1,c_2,\cdots,c_n|n\in N\}$ 生成 T 个聚类 $P=\{P_1,P_2,\cdots,P_T\}$;

/*行为计算阶段*/

$D_S=\varnothing$；$A\in P_A$；// P_A 是捕获到的已知行为

For each $P_i\in P(i=1,\cdots,T)$

Repeat:

Compute distance $d(P_i, P_A)$;

If $d(P_i, P_A)=0$，显示中间结果：P_i 不是隐形攻击行为;

Else if $d(P_i,P_A)=1$，隐形攻击行为 $D_S=P_i\cup D_S$;

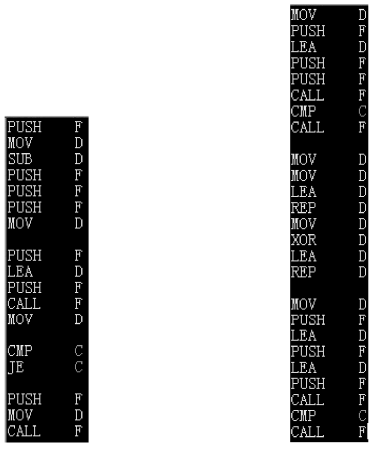
Until $i=T$;

返回隐形攻击行为集 D_S;

生成所有由基本块表示的行为指令序列 $B_1,B_2,\cdots$

End

大量实例分析表明，隐形攻击行为在基因指令的类型、顺序和执行频率上和协议正常行为有显著的不同。本文定义的指令序列 *IS* 特指标记过的指令序列，则 $IS=\{i_1,i_2,\cdots,i_k,\cdots\}$，其中，$i_k$ 是指第 k 条指令。为了便于进行聚类分析，每一条指令 i_k 被标记为 $l_k \in L=\{F,C,D\}$。本文假设一个协议二进制代码是由标记过的基因指令组成的，这些标记后的指令聚集成基本块，最终成为行为指令序列。如图 4 所示，根据标记后的行为指令序列，所有协议样本的行为可以自动聚类成行为聚类。



(a) 标记后的行业指令序列1　　(b) 标记后的行业指令序列2

图 4　标记后的行为指令序列实例

本文对协议样本集中标记后的二进制指令进行 *n*-gram 分析，生成的 *n*-gram 分布表示所有的行为。算法 1 的目标就是要感知和挖掘深藏在协议二进制代码中的隐形攻击行为。潜在隐形攻击行为的挖掘可以通过计算已知行为指令序列和新挖掘出的行为指令序列之间的行为距离来实现。假设 *A* 和 *B* 是本文已捕获的 2 个行为，且其功能已知，现在有一个行为未知的协议样本 *P*。*A* 和 *B* 形成 2 个行为模式 B_A 和 B_B，是因为行为 *A* 和行为 *B* 在指令序列上存在不同。为了研究协议 *P* 的行为，本文检查其指令序列的特征，将该特征和所有已知的行为模式进行对比，与哪一个行为模式最接近最匹配。如果该指令序列和行为模式 B_A 最匹配，其行为可表示为 *P.A*，这就意味着 $D(H(P.A),B_A)$ 小于本文预先设置的阈值，其中，*H* 表示 *n*-gram 分布，而 $D(\ )$ 则是行为距离函数。这说明未知协议 *P* 的行为 *P.A* 非常接近已知行为 *A*，它可以和 *A* 划归到同一个行为聚类中。如果已知行为 *A* 是正常的，则未知协议 *P* 的行为 *P.A* 也可以认为是正常的。同样道理，如果未知行为 *P.B* 和已知行为 *B* 非常接近，则它可以和 *B* 划归到同一个行为聚类中。如果未知行为 *P.N* 的 *n*-gram 指令特征和所有已知的行为模式均相距甚远，本文可以认为 *P.N* 是一个新的行为模式，并且产生一个新的行为聚类，而该行为是隐形攻击行为的概率非常大。

本文设计了一个通用的指令聚类框架，把动态分析和静态分析有机地融合在一起。行为模式是根据 3 类基因指令的频率、数量和分布计算获得的，而行为距离又是通过计算不同行为模式向量之间的差异来获得的。在这种方式下，*A* 是行为 *A* 的指令序列，*A* 的行为模式 B_A 是由 3 类基因指令的频率和分布组成的向量来表示的。例如，行为模式 $B_A=(freq_A(F,D,C),distrib_A(F,D,C))$，则行为距离 $D(A,B)=|B_A-B_B|=(B_A-B_B)^2$。指令聚类的过程如图 5 所示。

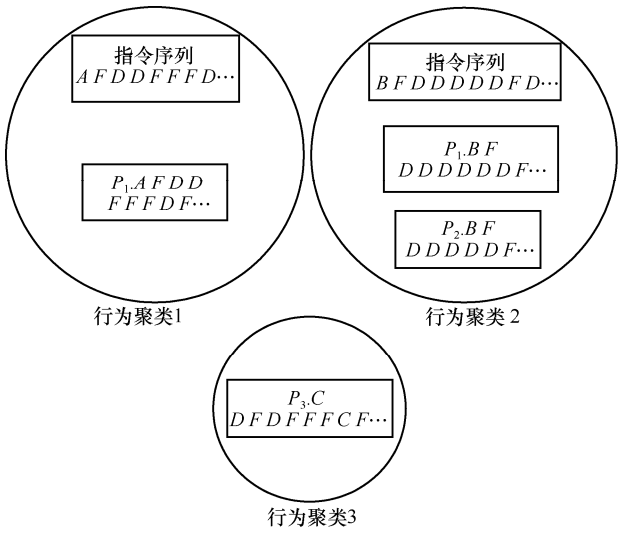


图 5　指令聚类过程实例

如图 5 所示，指令序列 *A* 和 *B* 是本文捕获到的 2 个已知行为，P_1、P_2 和 P_3 是 3 个未知协议的二进制程序。根据指令聚类算法，行为模式 B_A=(0.1,0.1,0.2,0.25,0.25,0.50)=0.435, 行为模式 $P_1.A$=(0.2,0.3,0.4,0.3,0.3,0.6)=0.83，则行为距离 $D(A,P_1.A)=|B_A-B_{P_1.A}|=0.395$，远远小于预先设定的阈值 3。实验结果表明挖掘出的行为 $P_1.A$ 和指令序列 *A* 非常相似，所以它们可以归并到一起形成行为聚类 1。同理，计算已知行为 *B* 和新挖掘出的行为 $P_1.B$ 和 $P_2.B$ 之间的行为距离，就可判断新行为和已知行为之间的差异度。行为模式 B_B=(0.40,0.30,0.30,0.40,0.20,0.40)=0.7，行为模式

$P_1.B$=(0.20,0.20,0.60,0.30,0.30,0.40)=0.782，行为模式 $P_2.B$=(0.25,0.25,0.50,0.25,0.25,0.50)=0.75。由此可知新挖掘出的行为 $P_1.B$ 和 $P_2.B$ 均和指令序列 B 相近，所以它们都可以归并到一起形成行为聚类 2。继续计算新挖掘行为 $P_3.C$ 的行为模式 $P_3.C$= (0.20, 0.40, 0.40, 1, 1, 2)=6.36，行为距离 $D(A,P_3.C)=|B_A-B_{P_3.C}|$=5.925 行为距离 $D(B,P_3.C)=|B_B-B_{P_3.C}|$= 5.66 均大于预先设定的阈值 3。显然，挖掘出的行为 $P_3.C$ 和行为 A 及行为 B 的差异均非常大，据此它单独形成了一个聚类，$P_3.C$ 就是一个潜在的隐形攻击行为。实例分析表明指令聚类算法可以感知并挖掘出潜在的隐形攻击行为，但该行为是否有害，以及它的具体功能是什么还需进一步验证执行。本文在自主研发的虚拟分析平台 HiddenDisc 上分析并验证了大量协议的隐形攻击行为，限于篇幅本文对此不再赘述，可参阅文献[10]。

4　实验及分析

4.1　实验结果

初始阶段，本文抽取了 25 个指令序列，代表 25 种不同的协议行为。为了更有效地实施聚类分析，本文捕获了公开协议的 25 个典型行为指令序列，如 HTTP、FTP、DNS 和 SMB 等协议。如对 HTTP 协议，本文捕获的指令序列是 Apache server 如何处理获取文件“index.html”的 HTTP GET request 以及由服务器产生的响应。对 FTP 协议，本文分析和抽取的指令序列是 FileZilla 服务器发送消息以及对连接的响应，同时还包括消息发送时，用户名和口令的接收行为。对 SMB 协议，本文分析和抽取到的行为是 Negotiate 协议请求接收开源服务器 Sambad 发送的消息。本文同样分析了一些经典的非正常行为指令序列，如口令嗅探、击键嗅探、后门访问以及 rootkiting 和 spying 等。不到 3 min 的时间，1 297 个未知协议样本依次在虚拟分析平台 HiddenDisc 上自动执行，根据指令聚类算法，所有行为距离都被计算出来。根据 25 个初始基本行为，挖掘出的所有行为形成了 193 个聚类。表 1 是其中 3 个行为聚类的分析实例。

如表 1 所示，行为 1 和行为 2 是在 HiddenDisc 平台上动态分析捕获到的指令序列，行为 3 没有被动态分析捕获到，却被指令聚类算法挖掘出来。指令聚类算法的实验结果显示这 3 个行为的基因指令特征分别为 B_1=($freq_1(F,D,C)$, $distrib_1(F,D,C)$)= (0.18, 0.10, 0.20,9,5,1)；B_2=($freq_2(F,D,C)$, $distrib_2(F,D,C)$)= (0.27,0.19,0.60,9,6,2)；B_3=($freq_3(F,D,C)$, $distrib_3(F,D,C)$)=(0.10,0.13,0.20,9,10,2)。行为距离分别是：$D(B_1,B_2)=|B_1-B_2|=(B_1-B_2)^2=2$；$D(B_1,B_3)=|B_1-B_3|=(B_1-B_3)^2=26$；$D(B_2,B_3)=|B_2-B_3|=(B_2-B_3)^2=16$。尽管这 3 个行为各不相同，但行为距离却揭示了深层秘密。大量实例研究显示，如果行为距离 $D(B_1,B_2)<3$，尽管 B_1 和 B_2 的功能不同，但是这 2 个行为的本质几乎相同，所以 B_1 和 B_2 可以归为同一个聚类。这就意味着如果 B_1 是正常行为，则 B_2 也是正常行为。若 B_1 为正常行为，而行为距离 $D(B_1,B_2)>3$，那么 B_2 就很有可能是隐形攻击行为。若行为距离 $D(B_1,B_2)>9$，根据实例分析的数据，则 B_2 就是典型的恶意行为，其指令序列必须进行进一步的分析。本文再次对这 3 个行为进行验证性执行，最终确定：行为 1 和行为 2 均是正常行为，尽管它们的功能不同，但行为 3 给操作系统目录中拷贝了大量可疑文件。通过指令聚类，本文从 1 297 个协议样本中挖掘出了 193 个潜在隐形攻击行为指令序列，最终在这 193 个指令序列中发现了 187 个恶意行为。

效率分析。从 1 297 个协议样本送入 HiddenDisc 分析平台到 193 个潜在隐形攻击行为全部挖掘出来共用时 173 s，其中，分析一个协议样本平均耗时仅为 0.13 s。目前，协议分析人员每天将面对数以千计的未知协议，而这样的分析效率是另人满意的。

根据本文提出的指令聚类算法，1 297 个协议样本最终被自动划分成 5 个行为聚类，其中 Cbot-3280 是一个拥有隐型攻击行为的样本，而 http-hidden 是本文开发的具有隐匿行为的协议。聚类分析结果如表 2 所示。

表 1　　**行为指令序列分析实例**

行为聚类	标记后的行为指令序列	是否隐形攻击行为
行为 1	*FDDFFFD FDFFD CC FDF FF DFDF D FFFDFF*	否
行为 2	*FDDFFFD DFFFD CC FDF FF DDDD CC FDFDFFD DFFFDF DFCC FDF D FFFDFF*	否
行为 3	*DFDFFFCF DDDDDDDD DFDFDFFCF*	是

表 2 **协议样本的行为聚类结果**

协议程序	公开：隐形基本块数	公开行为分布向量 $\boldsymbol{distrib}_{\text{pub}}(F_1,D_1,C_1)$	隐形行为分布向量 $\boldsymbol{distrib}_{\text{dorm}}(F_2,D_2,C_2)$	行为距离	运行安全性
聚类 1 (121)	5 200：110	(0.25,0.25,0.50)	(0.25,0.25,0.50)	0	安全
聚类 2(356)	4 430：100	(0.40,0.30,0.30)	(0.40,0.20,0.40)	0.02	安全
聚类 3(818)	4 690：90	(0.20,0.20,0.60)	(0.30,0.30,0.40)	0.06	安全
聚类 4(Cbot-3280)	2 750：1 330	(0.10,0.10,0.80)	(1.00,1.00 ,2.00)	5.34	不安全
聚类 5(http-hidden)	5 000：1 990	(0.20,0.40,0.40)	(1.70,1.10,2.20)	5.94	不安全

如表 2 所示，基因指令的不同分布意味着不同的协议行为。在聚类 1 的 121 个协议样本中，基因指令的分布完全相同，没有区别。尽管有少量的隐形行为被挖掘出来，但是从基因指令的分布来看，这些行为和公开行为相比没有本质区别。行为距离为 0 表明所有的行为都是完全公开透明的。因此，本文认为这个聚类的协议在运行上是安全的。在聚类 2 和聚类 3 的一些协议中本文发现了隐形行为，基因指令的分布也显示出一些不同。行为距离的计算结果表明，尽管公开行为和隐形行为之间存在一些差别，但特征距离却很小，远低于本文预先设定的阈值。验证执行发现这些所谓的隐形行为其实是一些保护代码，如加密和混淆等功能。因此，这 2 个聚类的协议在运行上也认为是安全的。然而，Cbot-3280 和 http-hidden 这 2 个协议则完全不同。首先，隐形行为的比例显著增加；其次，基因指令的分布也明显不同。行为距离的计算结果均超出了阈值，远高于前 3 个聚类，这也意味着隐形和公开行为之间的特征距离显著增加。因此，本文认为这 2 个协议的运行是不安全的。最终的验证性执行表明通过行为距离来决定协议的运行安全性是准确和可靠的。

4.2 讨论

未知协议隐形攻击行为的多样性和隐蔽性严重削弱了经典的基于特征检测方法的有效性。所以，本文并不是靠收集协议的特征信息来精确定义协议的正常行为模式，相反，本文尝试着去发现协议的正常和异常行为指令序列在距离上的变化规律。实验中用到的这 1 297 个协议样本均收集自 Internet,且每一个样本都在实验前手动分析过。研究关注协议二进制代码中 3 类基因指令的分布特征，本文重点跟踪和记录每一条指令序列当中基因指令的类型、数量和频率特征。通过对这 1 297 个协议行为的分析，本文发现协议的隐形攻击行为在基因指令的特征上和正常行为有较大的不同。众多协议行为最终形成了 5 类典型行为聚类，其基因指令的分布特点由图 6 所示。

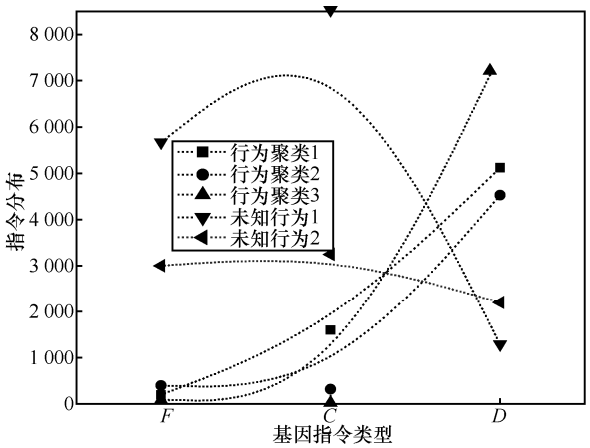


图 6 基因指令聚类实例

尽管这些协议的行为各不相同，但是从基因指令的分布上来看，规律还是明显的。前 3 组行为有一个共同特点，那就是函数调用相关指令很少，指令多集中在控制流转移和数据处理上。鉴于网络协议的主要行为就是发送和接收消息，所以，包含较多的数据处理指令属于正常现象。但是后 2 组行为的指令分布却有明显的不同。指令多集中在函数调用和控制流转移上，而数据处理指令却相对偏少。对于一个网络协议来说，它的主要功能不是数据处理，却执行大量的其他操作，那么这些行为显然不正常。大量实例研究表明，那些和正常行为距离接近的行为通常也是正常行为，而相反则很有可能是隐形攻击行为。协议的正常行为和隐形攻击行为可以通过指令聚类有效地区分出来。这样，验证性执行目标更明确，而且分析效率得到了极大的提升。验证性执行的结果表明，后 2 组未知行为均是隐形攻击行为，而且也均属于恶意行为。

对 HiddenDisc 平台方案和目前流行的分析机制进行了对比，特别是对隐形行为的识别能力进行了比较研究，比较的结果如表 3 所示。

由于无法获得上述分析工具，本文并不能使用这些方法分析自己捕获的协议样本，但是根据相关文献的描述，本文可以推断这些工具对于隐形行为的识别率。比较研究显示，经过不断努力和改进，已经有相当数量的工具和方法能很好地识别加密和混淆算法，而且识别率有了显著提高。然而这些方法对于隐形行为的识别却显得力不从心。尽管隐形行为通常被加密和混淆算法保护，但由于缺少对隐形行为特征的研究，即使发现了隐形行为的保护算法，仍然很难识别隐形行为。HiddenDisc 是目前所知能够挖掘和分析隐形行为最有效的工具，识别率在 80%以上，远高于已有的方法和工具。

需要强调的是，协议设计者可以通过代码重构、代码混淆等技术改变协议行为的表现形式，但却无法改变基因指令的特征，所以通过指令聚类来划分协议的行为是行之有效的方法。尽管指令聚类能够暴露出隐形攻击行为，但它却不能准确地确定该隐形攻击行为是否一定是恶意行为，因为还需要进一步验证执行，这就增加了分析的时间和成本。目前，本文是通过计算隐形攻击行为和正常行为指令序列的长度比例来评估协议的运行安全性，这就有更多的隐形攻击行为将导致协议运行越不安全。但是尽管有一些隐形攻击行为和正常行为差距较大，但它们也可能是良性行为，所以该评估方法也有可能造成一定的误判。本文对协议行为的研究尚处于初级阶段，这些规律是否适用于所有未知协议还不确定，仍需要继续深入研究。

5　结束语

本文提出了一种智能化的指令聚类方法，用于感知和挖掘大量未知协议中的隐形攻击行为。不同于以往的研究，本文的方案将动态污点分析和全新的指令聚类分析相结合，用来高效分析未知协议中的指令序列。在不到 3 min 的时间内，就完成了对 1 297 个协议样本的指令聚类分析，挖掘出 193 个隐形攻击行为，而验证执行的结果和手动分析结果完全一致，表明本文提出的方案在效率和准确性方面都是理想的，达到了预期的目标。

参考文献：

[1] CUI B, WANG F, HAO Y, et al. A taint based approach for automatic reverse engineering of gray-box file formats[J]. Soft Computing, 2015:1-16.

[2] BOSSERT G, GUIHÉRY F, HIET G. Towards automated protocol reverse engineering using semantic information[C]//Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security. 2014.

[3] NARAYAN J, SHUKLA S K, CLANCY T C. A survey of automatic protocol reverse engineering tools[J]. ACM Comput Surv, 2015, 48: 1-26.

[4] LI X D. A survey on methods of automatic protocol reverse engineering[C]//Proceedings of the 2011 Seventh International Conference on Computational Intelligence and Security. 2011: 685-689.

[5] CABALLERO D S J. Automatic protocol reverse-engineering: message format extraction and field semantics inference[J]. Computer Networks, 2012, 54(2): 451-474.

表 3　HiddenDisc 和主流解决方案的对比

方案名称	解决方案	分析对象	可否抵御加密/混淆技术	可否识别隐形行为	识别率
推断协议状态机[7]	监听网络消息流	网络消息	否	否	0
动态污点分析[1,11]	抽取协议程序处理输入数据的一些特征	协议程序	可	可	10%~20%
堡垒逆向工程[12]	识别堡垒中来自其他恶意软件的恶意代码片断及其开源应用	僵尸程序和网络消息	可	可	<50%
识别隐匿功能[13]	根据观察到的行为抽取和建模部分新的恶意软件二进制代码	恶意程序	否	可	20%~40%
恶意软件分类[14]	分析恶意软件的相似性来划分恶意软件类别	恶意程序	可	可	约 20%
探究多执行路径进行恶意软件分析[15]	识别仅当特定条件满足时执行的恶意动作	恶意程序	可	可	约 60%
僵尸网络组活动探测器[12]	使用 DNS 流量检测僵尸网络	僵尸网络消息(DNS 流量)	可	可	约 30%
HiddenDisc	使用指令聚类挖掘协议隐匿行为	协议程序和网络消息	可	可	≥80%

[6] JOÃO A N N. Automatically complementing protocol specifications from network traces[C]//European Workshop on Dependable Computer, 2011: 87-92.

[7] MENG F Z, LIU Y, ZHANG C R, et al. Inferring protocol state machine for binary communication protocol[C]//Advanced Research and Technology in Industry Applications (WARTIA). 2014:870-874.

[8] HAN K, LIM J H, IM E G. Malware analysis method using visualization of binary files[C]//Proceedings of the 2013 Research in Adaptive and Convergent Systems, Montreal, Quebec, Canada, 2013.

[9] 苏璞睿, 杨轶. 基于可执行文件静态分析的入侵检测模型[J]. 计算机学报, 2006, 29: 1572-1578.
SU P R, YANG Y. Intrusion detection model based on executable static analysis[J]. Chinese Journal of Computers, 2006, 29: 1572-1578.

[10] 胡燕京, 裴庆祺,庞辽军. 消息和指令分析相结合的网络协议异常行为分析[J]. 通信学报, 2015,36(11): 147-155.
HU Y J, PEI Q Q, PANG L J. Message combined with instruction analysis for network protocol's abnormal behavior[J]. Journal on Communications, 2015, 36(11): 147-155.

[11] LIN W, ZHU Y F, SHI X L. A method of multiple encryption and sectional encryption protocol reverse engineering[C]// 2014 Tenth International Conference on Computational Intelligence and Security (CIS). 2014: 420-424.

[12] RAHIMIAN A, ZIARATI R, PREDA S, et al. On the reverse engineering of the citadel botnet[J]. Foundations and Practice of Security, 2014:408-425.

[13] COMPARETTI P M, SALVANESCHI G, KIRDA E, et al. Identifying dormant functionality in malware programs[C]//IEEE Symposium on Security & Privacy, 2010: 61-76.

[14] KANG B, KIM T, KWON H, et al. Malware classification method via binary content comparison[C]//ACM Research in Applied Computation Symposium. 2012: 316-321.

[15] NATANI P, VIDYARTHI D. An overview of detection techniques for metamorphic malware[J]. Intelligent Computing, Networking, and Informatics. 2014:637-643.

作者简介：

胡燕京（1980-），男，陕西西安人，博士，武警工程大学讲师，主要研究方向为信息系统安全防护、网络协议逆向分析。

裴庆祺（1975-），男，广西玉林人，西安电子科技大学教授、博士生导师，主要研究方向为数字内容保护与无线网络安全。