摘要

随着万物互联时代的到来，由僵尸网络导致的 DDoS 攻击，勒索加密，钓鱼邮件，信息泄露等安全事件层出不穷。伴随大数据、物联网的发展，僵尸网络的危害涉及工控系统、IoT 设备、移动安全、云服务、电信服务等多个领域。同时，在国与国之间的网络对抗中，僵尸网络作为一种有效的网络武器能够对国家公共基础设施造成巨大的危害，在未来第五维网络空间战场中将发挥重要的作用。因此，研究僵尸网络的核心架构及关键技术，深入分析僵尸网络的行为特征及交互特点，掌握僵尸网络的进化方向及演变规律，对提升僵尸网络防护能力，保障网络空间安全有着重要的意义。

文章以僵尸网络的行为作为切入点，通过分析网络日志，识别僵尸网络在不同状态下的行为特征，结合行为时序关系，构建僵尸网络的行为识别模型。针对僵尸网络的不同行为，定义僵尸网络行为细粒度标签，结合机器学习相关算法，将散乱的行为标签加以关联，实现对僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的画像。主要研究内容包括：

（1）僵尸网络行为模型研究。通过分析僵尸网络的生命周期，建立僵尸网络行为标签与周期状态之间映射关系，研究僵尸网络在感染、维护、潜伏、传播、攻击、销毁六个状态下的转换行为及时序关系，提出基于 CRF（Conditional Random Field，条件随机场）的僵尸网络检测模型。

（2）僵尸网络行为标签研究。从标签提取，标签定义和标签识别三个方面论述了行为标签技术。以 Mirai 僵尸病毒为例，提取其行为特征，结合多种病毒分析结果归纳出僵尸病毒行为的特征元数据，然后根据行为标签的属性对其进行分层，设计标签结构，最后，根据实际情况使用多种方法识别僵尸网络行为标签。

（3）僵尸网络多维画像研究。研究基于行为分析僵尸网络画像体系结构，构建基于属性、规则和实例的僵尸网络本体。针对僵尸网络行为标签，结合机器学习的相关算法，将散乱的僵尸网络行为标签进行结构化关联，实现僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的画像。

（4）基于行为标签的僵尸网络画像仿真平台实现。以 ELK 日志平台作为基础架构，结合 Bro 入侵检测系统，Spark 实时计算系统以及 sklearn，tensorflow 等相关机器学习工具实现网络流量及日志的实时采集、存储、分析，实现僵尸网络画像。

关 键 字：僵尸网络， 画像， 行为标签， 机器学习

ABSTRACT

With the development of Internet industry. Security incidents such as DDoS attacks, ransomware encryption, phishing emails, and information leakage caused by botnets have emerged in an endless stream. With the development of big data and the Internet of Things, the dangers of botnets involved industrial control systems, IoT devices, mobile security, cloud services, and telecommunication services. At the same time, during the network confrontation between countries, botnets, as an effective cyber weapon, can cause great damage to the country’s public infrastructure and will play an important role in the future fifth-dimensional battlefield in cyberspaces. Therefore, research on the core architecture and key technologies of botnets, in-depth analysis of botnets' behavioral characteristics and interaction characteristics, the evolutionary direction and evolution of botnets are important, This research will improving the protection of botnets defense ability and ensuring the security of cyberspace.

This paper taking the behavior of the botnet as an entry point, we identify the behavioral characteristics of the botnet in different states by analyzing the network log, and construct the botnet's behavior recognition model by combining the behavioral timing relationships. Aiming at different kind of behaviors of botnets, we define a fine-grained label system for botnet network behavior. Then, the scattered behavior tags are associated by some machine learning algorithms. Finally, this paper presents a portrait of botnets in terms of survivability, offensive capabilities, spreading capabilities, identity characteristics, activity patterns, and local assets. The main contributions of this work can be summarized as follows:

(1) Botnet behavior model. We establish a mapping relationship, which between botnet behavior labels and cycle status by analyzing the life cycle of botnets, then, we study the transformation behaviors and timing relationships of botnets in the states of infection, maintenance, latency, infection, attack, and destruction. Finally, we come up with a model for botnet recognition which base on CRF.

(2) Botnet behavior label. Behavioral labelling technology includes three aspects: label extraction, label definition and label identification. We take Mirai zombies virus for example. Firstly, we summarize the feature metadata of zombie virus behavior by extracting and

analyzing their behavioral characteristics. Secondly, we layer it according to the attributes of the behavior label and design the label structure. Finally, according to the actual situation, we identify botnet behavior labels using a variety of methods.

(3) Research on multi-dimensional portraits of botnets. Research on the botnet portrait architecture based on behavior analysis. Build botnet ontologies based on three basic aspects: attributes, rules, and instances. Aiming at the botnet behavior label and combining the related algorithm of machine learning, the botnet behavior labels of the scattered bots are structured and related, and the botnet's survivability, attack ability, spreading ability, identity characteristics, activity pattern and local assets to achieve botnet portrait.

(4) Implementation of a botnet portrait platform based on behavioral labels. Using the ELK log platform as an infrastructure, combined with Bro intrusion detection systems, Spark real-time computing systems, and sklearn, tensorflow and other related machine learning tools to achieve real-time network flow log collection, storage, analysis, to achieve botnet portrait.

Keyword: botnet, portrait, behavior labels, machine learning

插图索引

表格索引

符号对照表

符号	符号名称
L_1	频繁一项集
L	频繁项集
D	标签集合
g_j	神经元梯度
w	神经元权值
θ	神经元偏向

缩略语对照表

缩略语	英文全称	中文对照
C&C	Command and Control Channel	命令控制信道
CRF	Conditional Random Field	条件随机场
DDoS	Distributed Denial of Service	分布式拒绝服务攻击
DGA	Domain Generation Algorithm	域名生成算法
DNS	Domain Name System	域名系统
HDFS	Hadoop Distributed Filesystem	Hadoop 分布式文件系统
HTTP	Hyper Text Transfer Protocol	超文本传输协议
ICMP	Internet Control Message Protocol	因特网控制消息协议
IRC	Internet Relay Chat	互联网中继聊天
IoT	Internet of Things	物联网
LSTM	Long Short-Term Memory	长短期记忆网络
PCA	Principal Component Analysis	主成分分析
TCP	Transmission Control Protocol	传输控制协议
TTL	Time To Live	存活时间
UDP	User Datagram Protocol	用户数据报协议

目录

第一章 绪论

1.1 研究背景和意义

随着计算机行业的不断发展，互联网与各个行业的融合持续加深，在万物互联的趋势之下，已经不存在能够独善其身的安全孤岛。国家公共基础设施，军队建设，金融资产，用户隐私等各个领域关键信息及服务均被暴露在互联网环境中，不得不面对纷繁复杂的网络安全问题。

一直以来，僵尸网络都是互联网环境中最具威胁的安全问题之一。僵尸网络指利用漏洞、后门、木马等渗透手段对网络空间中非合作个体进行远程控制，并使用一种或多种传播方式感染网络空间中其他个体，最终组建一个一对多的由控制者远程进行操控的网络攻击平台。僵尸网络为网络攻击提供了灵活的攻击手段和大量的攻击资源，具有数量庞大，高度可控，攻击灵活等优势，其危害包括：

（1）分布式拒绝服务攻击。DDoS 是当前网络空间中最难防御的攻击之一，DDoS 攻击使用了类似“闪电战”的思维，在极短的时间内，调动优势资源攻陷目标服务。僵尸主控机向僵尸网络中的所有僵尸主机发出指令，控制全部僵尸主机在特定的时间段同时访问目标主机，最大限度消耗目标主机网络资源，使其不能正常提供服务。由于僵尸网络规模庞大，协同攻击效率高，且结合协议特征，能够对访问流量进行成倍放大，并使用反射攻击等手段进行隐藏。因此，DDoS 攻击被认为是网络攻击中最有效，最难以抵御的网络攻击。

（2）海量垃圾邮件。僵尸主机通过设置代理进行垃圾邮件发送，在僵尸主控机的统一指挥下，僵尸网络可以传播大量的钓鱼邮件、色情信息、骚扰广告等。

（3）数据窃取。在组建僵尸网络的过程中，僵尸主控机已经获取了僵尸主机的部分权限，因此僵尸主控机能够轻易窃取僵尸主机的各类敏感信息。典型的窃取方式是键盘记录软件，通过识别键盘键入的关键字符，获取用户的关键信息，如银行账号密码，个人隐私等。同时，僵尸主控机也可以通过嗅探僵尸主机网卡获取敏感内容。

（4）舆情操纵。根据赛门铁克研究报告显示，2016 年，美国大选期间，全球范围内僵尸网络异常活跃。各方利益团体通过僵尸网络影响社会舆论甚至选票结果，达到其政治目的。由于僵尸网络规模庞大且高度可控，通常能够凭一己之力干扰舆情导向，使其成为了政治团体较量的工具之一。

（5）资源滥用。僵尸主控机在获取僵尸主机的部分权限后，即被赋予了调用其计算资源的能力。黑客能够通过操纵僵尸网络点击固定广告内容进行获利，或使用 ZeroAccess 调用僵尸主机计算资源进行比特币挖矿获利等。

随着互联网时代的到来，僵尸网络的规模空前强大，时刻威胁着网络空间安全。作为最具威胁网络安全问题，国内外各类安全组织及学者时刻关注的僵尸网络的发展态势。如图 1.1 及图 1.2 所示，CN/CERT 中国互联网 2016 年网络安全报告指出，2016 年我国僵尸主机数量最高达到 380 万，中国承受了全球约 84%的僵尸网络引发的 DDoS 攻击，属于僵尸网络重灾区[1]。赛门铁克公司在 2017 年第四季度全球范围内检测到的 12 万余起由僵尸网络发动的网络攻击中，中国遭受了 3 万余起，在 87 个受灾国家中排名第一[2]。

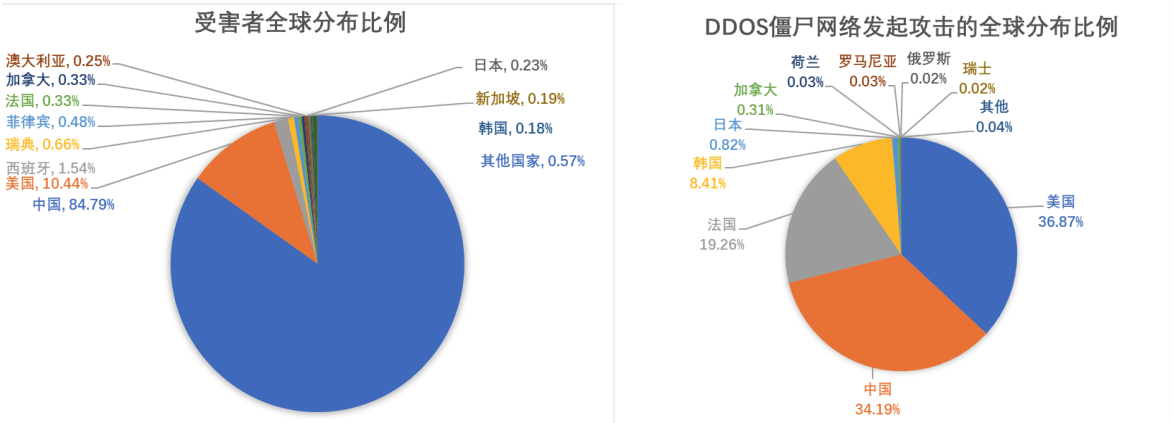
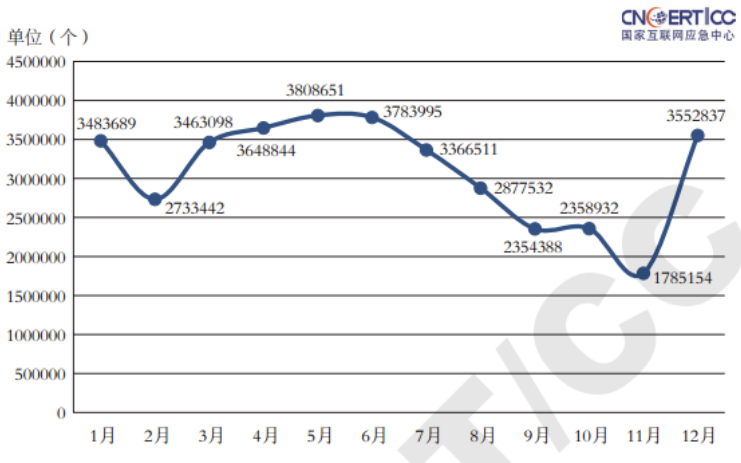


图 1. 1 全球发起 DDoS 攻击以及承受 DDoS 攻击分布比例



[1]

图 1. 2 2016 年木马或僵尸程序受控主机 IP 地址数量按月度统计

随着安全手段的不断加强，僵尸网络的攻击方式也在不断变化，二者在相互博弈中不断进化。进入二十一世纪以来，互联网发展的主题词包括大数据，云计算，人工智能，物联网等，僵尸网络无一不从中汲取新技术加以利用。

僵尸网络为各类攻击提供了平台，其攻击具有很强的可操作性和扩展性。在攻防

对抗中，僵尸网络通过引入各种先进技术对僵尸主控机进行隐藏。如 Zeus 中，使用 P2P 技术进行通信，使得 C&C 主机难以被发现，亦或使用 DGA 技术随机生成域名与 C&C 进行通信，以应对“黑名单”等安全机制[3]。

随着攻防技术的不断对抗，僵尸网络呈现出 APT（Advanced Persistence Threat 高级持续威胁）特性。如 2012 年的“燕子行动”，攻击者没有使用 PC 作为僵尸网络载体进行简单粗暴的 DDoS 攻击，而是大量使用性能更强、更稳定的 web 服务器进行攻击，而且在攻击过程中不断调整攻击方式以应对目标流量清洗策略，达到最大限度的消耗目标带宽，该次攻击持续一年之久，对美国多家银行造成了巨大的损失。

物联网技术的不断发展使得摄像头、打印机、路由器等网络设备均可以作为僵尸主机对网络空间构成威胁。2016 年 10 月 21 日，一款名为 Mirai 的 IoT 僵尸网络攻击域名提供商 DYN，从而导致了美国东海岸地区大面积网络瘫痪，包括 Amazon、Spotify、Twitter 等在内的多家互联网巨头均受到影响[4]。而 2016 年 9 月 20 日 Mirai 针对法国网站 OVH 的 DDoS 攻击流量达到了 1.5Tbps，打破了 DDoS 攻击强度的记录，其攻击能力足以使网络基础较为薄弱国家的网络设施完全陷入瘫痪。由此，IoT 型僵尸网络成为当前僵尸网络发展的趋势。

综上所述，僵尸网络作为经久不衰的网络空间主要威胁之一，其破坏性极大，且控制机制复杂，影响范围广泛，为构建网络空间安全体系带来了巨大的挑战。随着互联网的飞速发展，攻防之间的不断博弈使得僵尸网络在感染传播，通信隐藏，攻击手段，维护机制等多个方面均发生了改变。因此，为了应对僵尸网络的威胁，研究僵尸网络行为特征，设计僵尸网络检测模型，刻画僵尸多维特征，具有重要的现实意义。

1.2 国内外现状

互联网的飞速发展为僵尸网络的不断进化提供新的技术支持，随着国与国之间的网络对抗不断加剧，僵尸网络成为第五维网络空间战场的有力武器，众多的安全企业、研究院所、军队机构投入了大量的人力物力对其进行研究。

近年来，对僵尸网络的研究主要集中在僵尸网络传播机制、寻址机制、交互协议、隐匿机制、检测识别、拓扑机构、感染目的等方面，并取得了一定的成果。

规模庞大为僵尸网络的主要攻击优势，为了保障僵尸网络的攻击效果，僵尸控制者需要尽可能的感染和控制更多的僵尸主机。Sheharbano 等人[5]将僵尸网络的感染机制分为主动感染和被动感染，主动感染机制下，僵尸网络能够在没有人为操纵的情况下进行脆弱主机识别和感染，类似蠕虫病毒的主动感染，僵尸主机在对内网或外网扫描过程中，发现设备漏洞并进行攻击载荷投递，完成设备的感染[6]；被动感染机制下，需要人为的进行干预，使用网络钓鱼、口令猜测、共享介质等社会工程方法，利用人

性漏洞进行感染，如 Stuxnet 使用 APT 攻击的方法入侵物理隔离的伊朗核设施[7]，典型的网络钓鱼如 ZeroAccess[8]，将自身伪装成游戏插件进行传播。Aymen 等人[9]提出在 IRC 环境中，使用基于 IDS 及多阶段响应机制检测受感染的 IRC 僵尸网络成员。Carl 等人[10]提出使用僵尸网络的流量特征识别僵尸主机，通过对比正常主机流量与僵尸主机流量，提取其不同特征，结合机器学习的方法，识别僵尸主机流量，具有极高的准确性。Sushil 等人[11]使用朴素贝叶斯算法解决僵尸网络识别问题，提出在不对网络流量包进行深度解析的前提下，通过对网络行为的识别。使用朴素贝叶斯对行为进行关联分析，从而识别僵尸网络流量，该方法大大减小了安全策略带来的耗时操作。伦敦大学的 Stefano 等人[12]根据当前僵尸网络大多使用 DGA 算法随机生成域名与 C&C 主机进行交互这一特点，开发了一套智能识别系统：Phoenix。该系统通过对已知 DGA 算法的学习，能够分辨出不同的 DGA 种群，识别出大量域名中隐藏的 C&C 域名，为僵尸网络的阻断提供依据。Endgame 公司的 Jonathan 等人[13]提出使用循环神经网络的 LSTM 解决 DGA 域名识别问题，通过对 32 种 DGA 机制的研究，使得对僵尸网络域名的识别率达到了 99%。Sherif 等人[14]针对 P2P 僵尸网络 C&C 主机频繁更换，难以检测的问题，在路由节点进行僵尸网络 P2P 流量的识别，结合僵尸网络攻击前频繁通信的特征，识别僵尸网络 C&C。FeiFei Li 等人[15]提出使用 Deeplog 的方法，通过对日志进行解析，识别出其中的网络行为，并结合时序特征，建立了隐马尔可夫模型以识别僵尸网络。

在僵尸网络研究方面，国内的科研人员做出的贡献同样不可忽视。早期诸葛建伟等人[16]就僵尸网络进行研究，列举了僵尸网络对抗的关键技术及现实问题。方滨兴等人[17]提出一种改进型的 C&C 信道识别模型，通过识别僵尸网络中的异常指令来检测僵尸网络。清华大学吴建平等人[18]提出新的 P2P 僵尸网络体系架构，结合互联网发展的新特性，将其定义为构架机制，命令机制，控制机制，攻击机制，生存机制等五大核心问题，为进一步研究僵尸网络对抗关键技术提供了指导。国防科学技术大学的贾焰等人[19]针对在大数据环境中，僵尸网络 C&C 控制命令难以识别的问题，提出周期性识别僵尸网络模型，结合时间窗口识别僵尸网络的方法。北京邮电大学李可等人[20]针对僵尸网络隐匿机制越来越完善的现状，提出通过识别僵尸网络的行为，构建僵尸网络模型，并对其进行分析的方法。

综上所述，随着互联网技术的不断更新，僵尸网络的相关技术也在不断进化，大数据、云计算、物联网等技术的发展，赋予了僵尸网络新的威胁能力，IoT 僵尸网络的兴起重新定义了网络武器的概念，云计算的托管机制为僵尸网络提供了更广阔的入侵目标，大数据技术的发展使得僵尸网络通信越来越难以发现。为了应对僵尸网络的威胁，安全从业者不能出现一刻放松，需要在不断的对抗中提升安全防御机制和抵抗攻击能力。

1.3 研究目标及研究内容

文章针对僵尸网络的传播机制、命令控制协议、寻址机制、攻击机制、隐匿机制，提出基于行为的僵尸网络画像技术。通过提取和分析僵尸网络日志中的特定行为元数据，定义僵尸网络行为细粒度标签，结合机器学习相关算法，将散乱的僵尸网络行为标签加以关联，实现僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的画像。

主要研究内容包括：

（1）僵尸网络识别模型研究。通过分析僵尸网络的生命周期，研究僵尸网络在感染态，潜伏态，攻击态，交互态，寻址态等多个状态下的行为转换时序关系，提出基于 CRF（Conditional Random Field，条件随机场）[21]的僵尸网络检测模型。

（2）僵尸网络行为标签研究。通过捕获 Mirai 僵尸病毒行为，提取其特征，结合多种僵尸病毒行为分析结果，归纳了僵尸网络的行为元数据。从标签提取，标签定义和标签识别三个方面论述了僵尸网络行为标签技术。根据行为标签的属性对其进行分层，设计标签结构，并根据实际情况使用多种方法识别僵尸网络行为标签。

（3）僵尸网络多维画像研究。结合机器学习的相关算法，将散乱的僵尸网络行为标签进行结构化关联分析，构建僵尸网络本体。完成僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度进行画像。

（4）基于行为标签的僵尸网络画像仿真平台实现。以 ELK 日志平台作为基础架构，结合 Bro 入侵检测系统，Spark 实时计算系统以及 sklearn，tensorflow 等相关机器学习工具实现网络流量计日志的实时采集、存储、分析，实现僵尸网络画像。

文章的主要贡献包括：

（1）僵尸网络识别模型。针对僵尸网络控制周期长，活动频繁的特点，提出基于 CRF 的识别模型，该模型能够很好的结合时序关系，将僵尸网络行为时序关系作为识别僵尸网络的重要指标，同时避免了局部最优问题，大大提高了识别准确率。

（2）僵尸网络多维画像。僵尸网络多维画像能够更准确、系统的描述僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的特征，使对其分析达到了更深的层次，对于阻断僵尸网络，预测攻击意图，评估僵尸病毒能力，溯源和取证有着重要的意义。

1.4 主要工作及章节安排

文章共分为 8 章，其组织结构图如图 1.3 所示：

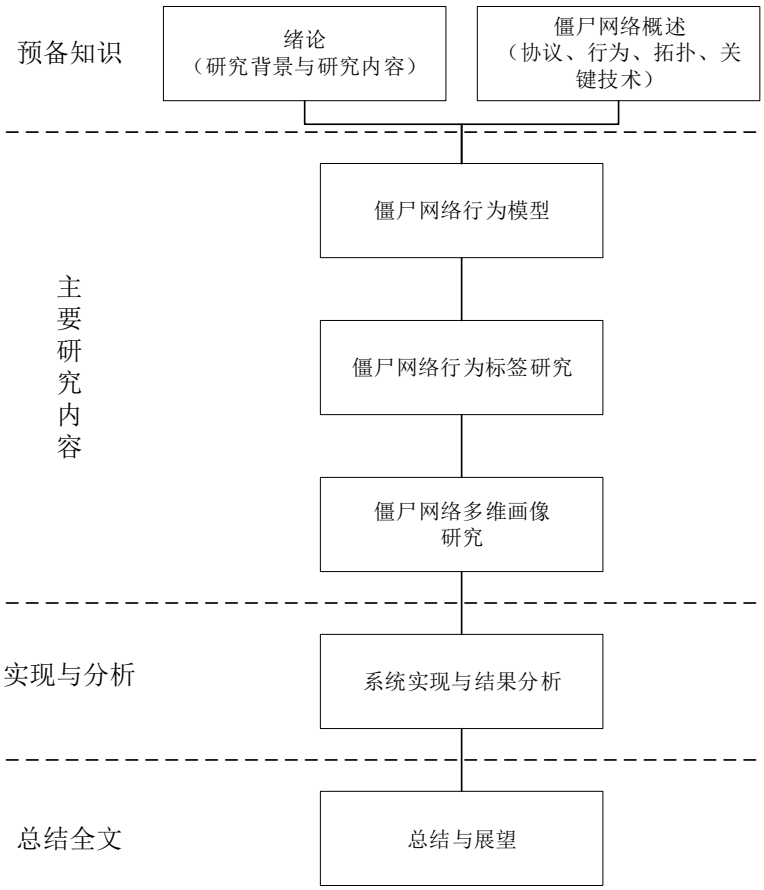


图 1. 3 论文组织结构

第一章：绪论部分，主要介绍了文章的研究背景及研究意义，介绍了僵尸网络的主要危害及课题面临的关键问题；论述了国内外科研机构对僵尸网络的主要研究成果，然后对文章的研究目标，研究内容及创新点进行了简要说明。

第二章：僵尸网络概述，主要对僵尸网络的基本概念，行为特征，运行机理进行综述性的描述，其中包括了僵尸网络的定义，范畴，行为，命令控制协议等，最终对僵尸网络与安全机制对抗关键问题进行总结。

第三章：僵尸网络行为模型，主要论述了基于行为标签的僵尸网络行为模型，并结合僵尸网络行为与状态之间的转换关系以及状态序列关系，提出基于 CRF 的僵尸网络行为检测模型。

第四章：僵尸网络行为标签研究，以 Mirai 这一典型僵尸病毒为例，提取其行为特征，结合多种病毒分析结果以及僵尸网络行为分析的要求，归纳出描述僵尸病毒行为的特征元数据，然后，根据行为标签获取的方式不同，将其分为基于元数据特征行为标签、统计型标签和关联型标签，设计行为标签结构，最后，结合行为的实际情况，采用多种方法对僵尸网络行为标签进行识别。

第五章：僵尸网络多维画像研究，首先构建基于属性、规则和实例的僵尸网络本

体。将对僵尸网络的画像归纳为生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度。然后，使用有监督与无监督的机器学习方法，将僵尸网络行为标签映射到不同的维度中，使用归一化的指标对僵尸网络的画像进行评估。

第六章：系统实现与结果分析，主要描述了整个基于行为标签的僵尸网络画像技术的总体架构和各个模块。其次，介绍起实验平台和相关技术。最后，对实验结果进行总结，对其性能进行评估。

第七章：总结与展望，主要对全文所有的工作进行总结，提出现有工作的不足之处以及未来的研究方向。

第二章　僵尸网络概述

2.1 僵尸网络基本概念

2.1.1 僵尸网络定义

僵尸网络（Botnet，Botnet 是 Robot Network 合体词）是指僵尸主控机（Botmaster）使用僵尸病毒对僵尸主机（Zombie）进行感染，并通过多种手段进行传播，形成的由僵尸主控机进行控制的大型网络。僵尸主控机通过命令控制信道（C&C，Command and Control Channel）对僵尸网络进行一对多的控制，对其下发攻击、维护、感染、信息回传等指令，僵尸主控机并不直接登录僵尸主机，攻击指令下发后，僵尸主控机通常会断开与 C&C 的连接，控制命令能够在僵尸网络中自动传播和执行。通常情况下，僵尸网络的数量庞大，分布遍布全球，其攻击手段多样，具有高度的可控性与协同性，因此，极难防御和溯源。其基本组成如图 2.1 所示。

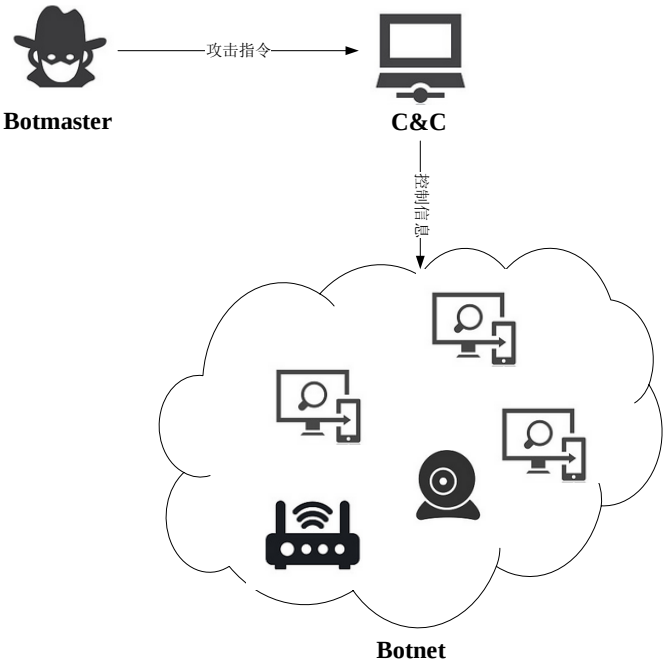


图 2. 1 僵尸网络基本组成

图 2.1 为集中控制式的僵尸网络，是僵尸网络比较典型的一种，同时，还存在着 P2P 僵尸网络。根据图 2.1，将僵尸网络做出如下定义：

$$Botnet = (Botmaster, Zombie, C\&C)$$

僵尸主控机，僵尸主机和命令控制信道构成了僵尸网络的三元组。僵尸主控机是整个僵尸网络的实际控制者，发出攻击，维护，感染，信息回传等控制消息，协同整个僵尸网络发出动作行为；命令控制信道是整个僵尸网络构建的核心，定义了僵尸网络的交互协议，拓扑结构，控制指令和网络资源；僵尸主机是僵尸网络攻击的直接实施者，随着物联网的发展，僵尸主机已经不局限于个人电脑和服务器，网络摄像头、路由器等 IoT 设备也成为僵尸网络中意的感染对象。

2.1.2 僵尸网络范畴

僵尸网络本质上是融合多种攻击手段的协同攻击平台，其攻击手法灵活多样，尤其是在感染传播阶段，其攻击渗透方式与蠕虫、木马、后门等极其类似。然而，僵尸网络也拥有其自身的特性，僵尸网络与其他恶意软件对比如表 2.1 所示：

表 2. 1 常见恶意代码对比

代码类别	传播性	可控性	窃密性	协同性	危害性
僵尸网络	可控传播	可控	有	有	全部控制：高
蠕虫病毒	主动传播	不可控	无	无	网络流量：中
木马程序	无	可控	有	无	全部控制：高
后门软件	无	可控	有	无	完全控制：高
勒索病毒	可控传播	不可控	无	无	感染文件：中
间谍软件	无	不可控	有	无	信息泄露：中

如表 2.1 所示，蠕虫（Worm）病毒[22]是一段能够通过计算机网络独立进行传播的恶意代码，一旦被感染，即通过网络主动传播，根据程序中定义的攻击行为对感染主机进行破坏，不受攻击者的控制。而僵尸网络的感染过程是高度可控的。

木马（Trojan）程序[23]的运作方式与僵尸网络极为相似，均具有高度可控性，攻击方通过多个跳板隐藏自身，对被攻陷主机进行控制，窃取信息或发动 DDoS 攻击等。然而，僵尸网络通常是数量庞大的僵尸主机的协同工作，木马通常指针对单个主机进行；同时，控制者在操作木马时，需要持续在线，增加了暴露风险，而僵尸网络通过 C&C 控制整个网络，攻击指令发出与攻击事实发生在时间上不同步，能够更好的隐藏僵尸主控机。

间谍软件（Spyware）[24]通常用于商业竞争及国家对抗中，面向相对封闭的内网环境，一般不可控，旨在对方不察觉的前提下窃取信息，不需要多个节点协同工作。而僵尸网络在内网和外网环境中均适用，且被感染成为僵尸主机后，窃密功能是可选的。

勒索软件（Ransomware）[25]通过获取计算机权限后对其文件等关键信息加密，从而进行勒索。勒索软件拥有可控传播性，在形式上与僵尸网络类似，均组成了一个被感染的网络，然而，僵尸网络在后续会进行协同工作，而勒索软件在攻陷受控主机后便不再与其进行交互，只需等待获取勒索到的电子货币后将密钥发回即可。

2.2 僵尸网络行为

文章的研究对象为僵尸主机，如没有特殊说明，文章所指僵尸网络行为均为僵尸主机的行为。在与其他网络攻击手段区分过程中，僵尸网络呈现出感染传播，规模庞大，高度可控，隐匿机制健全等特性。僵尸网络为了保证其攻击效果，在这些方面均表现出其独有的行为特征。则僵尸网络的典型行为如图 2.2 所示，主要包括传播、寻址、潜伏、交互及攻击。

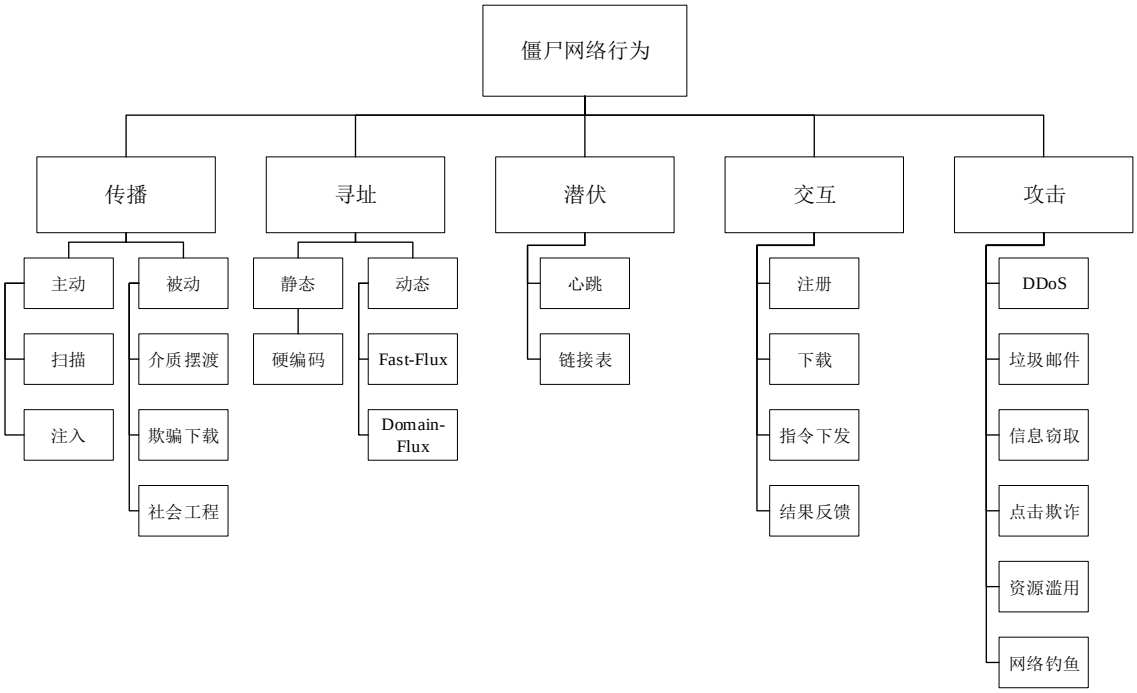


图 2. 2 僵尸网络行为

（1）传播

僵尸网络发动的攻击之所以难以防御，主要因为庞大的规模为其提供了大量的网络资源，在网络空间的消耗战中，哪一方拥有更多的资源，更好的策略，哪一方就能获胜。因此为了保证僵尸网络的攻击效率，其前期主要目标是不断扩充僵尸主机数量。根据是否需要人为介入，将僵尸网络的传播机制分为主动式和被动式。

（a）主动传播机制

僵尸主机能够在无需人为介入的情况下自主扫描内网及外网，发现漏洞并加以利用。主动传播机制通常包括扫描和注入两个动作。扫描过程中，僵尸主机对其内网或外网进行扫描和针对性的嗅探，发现网络中存在可利用漏洞的主机作为潜在的攻击目标，并利用漏洞进行远程登录或权限提升。在获取部分权限后，进行第二个动作：注入，将僵尸程序直接下载安装至受害主机中，受害主机被感染后，根据僵尸病毒的指令，自动对网段中的网络设备进行扫描和嗅探，重复传播的过程，不断扩大僵尸网络规模。有些僵尸网络直接使用蠕虫病毒进行传播。如 Storm 和 Sinit[26]使用了类似蠕虫的病毒进行传播。

（b）被动传播机制

被动传播机制在一定程度上需要人为干预，主要包括如下形式：

介质摆渡。该方法广泛的使用 U 盘作为载体进行传播，对工业控制系统等物理隔离系统进行传播和控制，Stuxnet 是一种具有高针对性的 APT 攻击手段，其正是通过 U 盘摆渡的方式入侵了伊朗布尔市核电站的网络系统[27]，造成设备故障，延缓了伊朗核进度。

欺骗下载。黑客通过伪造主流网站信息，或者入侵网站的方法，在主页上精心布置欺骗下载项诱导用户下载僵尸病毒程序，而这类程序通常使用 JavaScript 或 ActiveX 编写，一经下载，便能够自动进行安装和执行。

社会工程。网络系统中最大的漏洞是人。利用人类的本能，如好奇心理，猎艳心理，贪婪等，巧妙地布置网络陷阱，诱导受害者主动下载僵尸程序，相比于上述两种传播方式，社会工程的感染方式更灵活，危害程度更大且难以防范。常见的社会工程方式包括“鱼叉”和“水坑”[28]，将恶意程序链接在钓鱼邮件中，诱导受害者主动下载和执行。

（2）寻址

寻址是僵尸网络构建的关键步骤，也是僵尸网络与防御体系对抗的主要战场。寻址过程指僵尸网络通过域名或 IP 与僵尸命令与控制服务器取得联系的过程，也称为“上线”。

僵尸网络的高度可控性保证了其攻击效率，高效的寻址方式能够构建更灵活可控的僵尸网络。寻址方式主要包括静态和动态两种。静态寻址方式指将 C&C 主机的 IP、域名、链接表、URL 等信息以硬编码的方式写入僵尸程序，这种方式产生的通信流量较小，通常不容易发现，缺点是安全人员通过代码调试和分析可以轻易得到获取，进而对僵尸主机进行阻断，甚至关闭 C&C 主机，使整个僵尸网络陷入瘫痪。动态寻址通过各种混淆方法，阻碍安全人员获取 C&C 主机的域名及 IP，常见的有 Fast-Flux 技术和 Domain-Flux 技术[29]。前者通过构建域名与地址的一对多关系，使用代理节点完成僵尸主控机与僵尸主机的通信，隐藏了 C&C 主机的真实 IP；后者通过在僵尸主控

机与僵尸主机中构建相同域名生成算法，使用特定的信息生成大量域名，僵尸主机通过不断请求随机域名与 C&C 主机建立连接。

（3）潜伏

僵尸网络的感染与攻击在时间上是割裂的。僵尸病毒在感染受害主机后，并不会立即发动攻击，而是有相当长的一段潜伏时间以等待僵尸网络规模增大，在潜伏期间，僵尸主控机不会频繁的发布控制指令。

僵尸网络在潜伏期间并不会表现出明显的恶意行为，但是需要周期性的保持与 C&C 主机的连接，即“心跳”连接。潜伏期间最显著的行为特征是定期的特殊报文交互。如在 IRC 型僵尸网络中，C&C 主机与僵尸主机通过 Ping/Pong 关键字保持通信[30]，Mirai 僵尸网络通过发送内容为“0000”的报文进行心跳验证[31]。

（4）交互

僵尸主机上线后，在僵尸病毒销毁前，会与 C&C 主机进行一系列的交互活动。主要包括：

（a）僵尸主机注册。僵尸主机上线后，僵尸主控机会对僵尸主机的基本信息进行注册，主要包括僵尸主机的操作系统、MAC 地址、僵尸 ID、网络资源等。

（b）资源下载。僵尸主控机为了保证僵尸网络的安全性和可控性，会不定时的更新僵尸网络配置及僵尸网络程序，或是在攻击前发布工具，僵尸主机需要根据交互地址主动获取这些资源并运行。

（c）指令下发。僵尸主控机通过 C&C 主机发布指令，操纵僵尸主机做出反应。

（d）结果反馈。僵尸网络将自身收集到的敏感信息，攻击指令执行结果等信息通过 C&C 主机反馈给僵尸主控机。

（5）攻击

攻击是最能够识别僵尸主机行为的特征。在黑产肆虐的今天，僵尸网络更多的用以牟利。因此，所有能够用获取利润的攻击方式，黑客们都会进行尝试，僵尸网络主要的攻击方式是分布式拒绝服务攻击。根据 DDoS 攻击种类的不同，其行为亦呈现出不同的特征，如使用小包方式进行 UDP flood 攻击，在流量上呈现出大量的针对某一特定 IP 的 64Byte 大小的数据包请求；使用 DNS 反射攻击，会发现大量的冒用某一固定 IP 的 DNS 查询信息，或是大量的 DNS 查询信息无返回的情况。同时，也包括敏感信息窃取。使用键盘记录、网络嗅探等方法，类似于 APT 攻击的方法，通常使用一系列 0day 进行组合，极难发现。除此之外，还有网络钓鱼，垃圾邮件，比特币挖矿[32]，点击欺诈等行为。

2.3 僵尸网络命令控制协议

僵尸网络的命令控制协议，即僵尸主机与命令控制服务器的交互协议是研究僵尸网络的核心。对于黑客而言，高效的、隐蔽的、健壮的命令控制协议能够大大的延长僵尸网络生命周期，提升僵尸网络攻击效率，进而提高僵尸网络产业利润，其失效将直接导致僵尸主控机对僵尸网络的控制权丧失，僵尸网络陷入瘫痪，分解为离散的受感染的孤立节点，无法阻止有效地网络攻击，其威胁亦大大降低。对于安全人员而言，分析僵尸网络的命令控制协议是研究僵尸网络的前提，准确的协议识别和深入分析能够有效地阻断僵尸网络与 C&C 的连接，定位 C&C 服务器，为瓦解整个僵尸网络提供依据。

2.3.1 IRC 协议

IRC 协议是僵尸网络最早使用的交互协议。其最显著的特点是构造简单和交互性好。攻击者可以通过任意服务器创建 IRC 服务控制僵尸网络，且 IRC 的交互是实时的，攻击者能够根据网络态势实时控制僵尸网络做出反应[33]。IRC 的拓扑结构如图 2.3 所示：

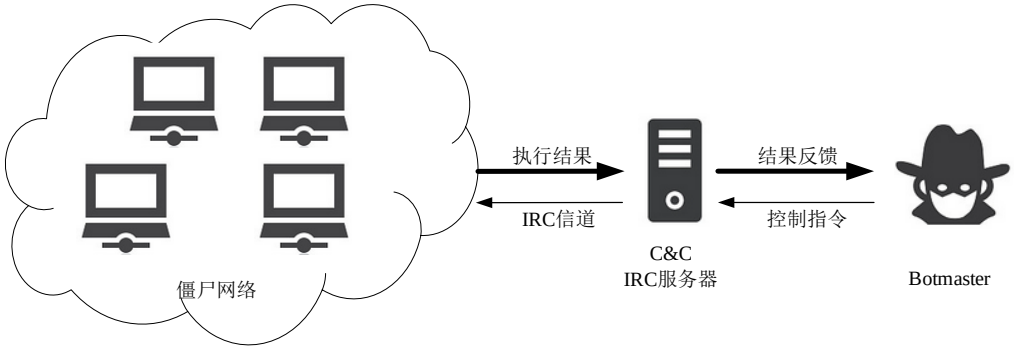


图 2. 3 IRC 型僵尸网络

IRC 型僵尸网络的构建步骤：首先构建僵尸网络与 C&C 服务器之间的通信。IRC 类型的僵尸网络会将 C&C 服务器的域名及端口号以硬编码的方式写入僵尸病毒，当僵尸病毒执行后，通过三次握手连接建立与 C&C 服务器之间的通信。此时，僵尸程序发送 NICK 指令加入 IRC 频道，进入 PING/PONG 状态等待僵尸主控机的指令。其次，构建僵尸主控机与 C&C 服务器间通信。僵尸主控机通过将预先设置在僵尸病毒中的控制密码发送至 C&C 主机，僵尸主机将密码与硬编码中的信息进行对比，完成僵尸主控机的认证。最后，整个僵尸网络构建完成。

IRC 型的僵尸网络构造过于简单，存在很多不足，主要包括：

（1）安全机制不健全，IRC 型僵尸网络通常需要将 C&C 服务器的域名及端口等

信息通过硬编码的方式写入僵尸程序中，安全人员很容易获取 C&C 地址并加以阻断。在僵尸网络进行攻击时，需要僵尸主机与僵尸主控机同时登陆，增加了僵尸主控机暴露的风险。

（2）认证机制不健全，安全人员通过逆向工程，很容易可以将自身伪装成僵尸网络加入 IRC 聊天室中，整个僵尸网络的信息随即暴露，部分黑客使用挑战应答的方式识别自己的僵尸主机，然而效果并不理想。

2.3.2 HTTP 协议

随着 web 应用的广泛使用，由于 HTTP 流量能够穿透防火墙和 IDS，使得 HTTP 协议成为僵尸网络的重灾区。僵尸网络将自身的通信流量隐藏在大量合法的 web 流量中，HTTP 型僵尸网络成为主流的 C&C 协议[34]，HTTP 型僵尸网络拓扑如图 2.4 所示：

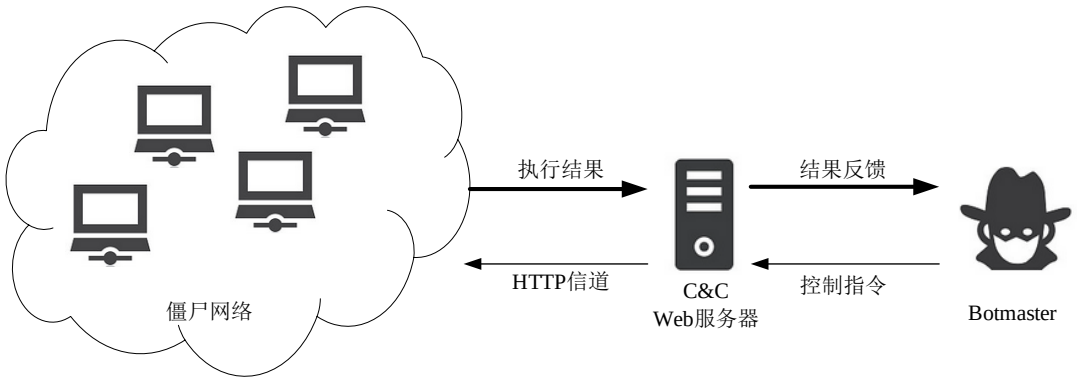


图 2. 4 HTTP 型僵尸网络

HTTP 型僵尸网络的构建步骤：首先，僵尸主控机将经过加密的控制指令发布在指定的 web 服务页面上，并将资源发布在指定的 url 链接中，等待僵尸主机主动获取；其次，僵尸主机根据僵尸病毒内置的轮训机制，周期解析内置硬编码的域名，访问并获取僵尸主控机发布的指令及资源；最后，僵尸网络将自身搜集的敏感信息上传至指定的 web 服务器，等待僵尸主控机回收。

HTTP 型僵尸网络使用 HTTP 协议进行通信，极大地隐藏了僵尸网络的通信流量，且僵尸主控机不需要与僵尸主机同时在线即可完成网络攻击行为，大大降低了僵尸主控机暴露的风险。然而，HTTP 僵尸网络与 IRC 僵尸网络类似，均使用了集中管控的方式，且需要将 C&C 服务器的 IP，域名等信息以硬编码的形式写入僵尸程序中，加大了 C&C 服务器暴露的风险，且抗打击能力不强。

2.3.3 Fast-Flux 协议

为了防止安全人员通过逆向工程定位 C&C 服务器，僵尸网络使用 Fast-Flux 技术隐藏 C&C 的 IP 地址[35]，提升僵尸网络的健壮性。其命令控制流程如图 2.5 所示：

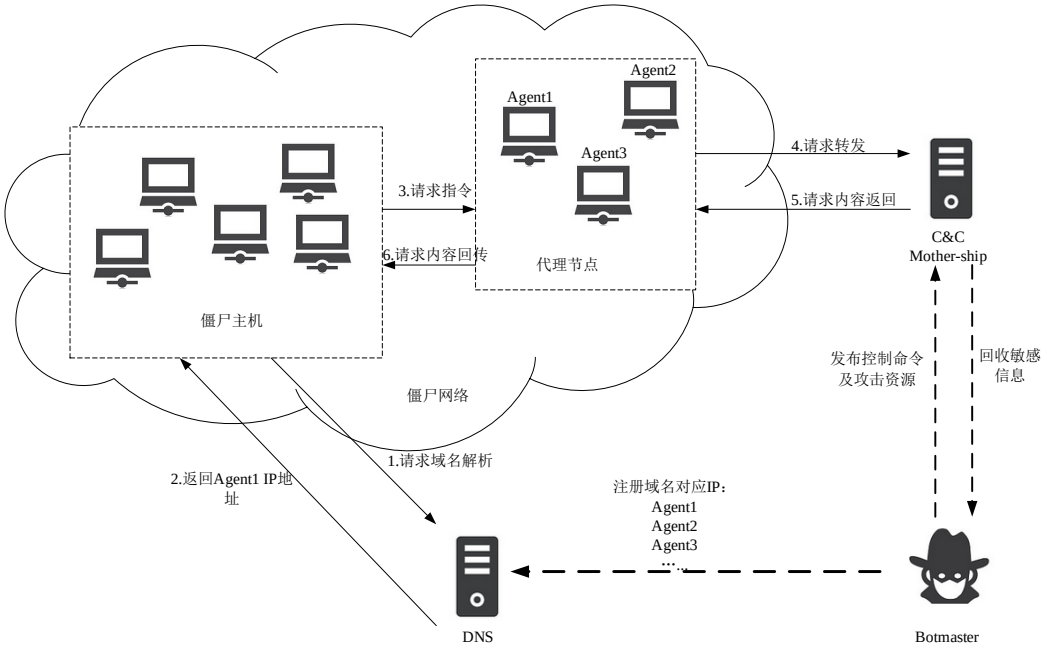


图 2. 5 Fast-Flux 命令控制流程

Fast-Flux 技术通过构建域名与 IP 的一对多关系实现 C&C 服务器的隐藏，如图 2.5 所示，当僵尸主机根据硬编码解析域名时，返回的并不是 C&C 服务器 IP 地址，黑客在权威 DNS 服务器并不直接注册 C&C 服务器的域名与 IP 的映射关系，而是注册一串代理节点（通常为僵尸主机）与僵尸程序硬编码域名之间的映射关系，僵尸主机将指令请求发送至代理节点，将代理节点作为转发站点协调僵尸主机与 C&C 服务器之间的通信，僵尸主控机只需要与 C&C 服务器通信即可。

一般情况下 DNS 服务器注册代价较高，有些黑客组织会自己部署底层域名服务器，以更灵活的建立 Fast-Flux 协议中域名与 IP 的一对多映射关系。主要的技术手段包括：Single-Flux 和 Double-Flux。Single-Flux 中，控制者为了更频繁的变更域名对应的 IP 地址，需要提供底层域名服务器，灵活的变换僵尸网络与 C&C 服务器的连接地址。Double-Flux 中增加了对 C&C 域名与 IP 之间的一对多映射关系。控制者会部署多个解析 C&C 服务器域名，不断修改顶级域名服务器中底层域名服务器的地址以隐藏 C&C 服务器。Fast-Flux 最初的设计愿景是通过建立域名与 IP 的一对多映射关系缓解服务器压力，当访问量超过服务器运行负荷时，进行访问流量分流。然而，在僵尸网络的 Fast-Flux 协议中，黑客只需要维护 C&C 服务器以及 DNS 服务器，僵尸

主机只能够建立与代理节点之间的连接，在溯源过程中，僵尸主控机往往在一系列跳板之后，很难追踪。当前环境下，对 Fast-Flux 僵尸网络的检测需要对 DNS 查询服务进行深度解析和评估，这类方法通常需要耗费大量的人力物力，且准确率并不高。

然而，Fast-Flux 协议还是表现出了不同于一般僵尸网络的行为特征，其 DNS 请求的生存周期短，连接 IP 地址分布广泛，这些行为均可以作为识别 Fast-Flux 僵尸网络的特征。同时，安全人员仍然可以通过设置域名黑名单达到僵尸主机下线的效果。

2.3.4 Domain-Flux 协议

Domain-Flux 协议中[36]，以硬编码方式写入僵尸程序的不再直接是 C&C 主机的域名，而是通过 DGA（Domain Generation Algorithm 域名生成算法）算法动态生成的一系列域名集合，真正的 C&C 主机域名隐含在其中。其命令控制流程如图 2.6 所示：

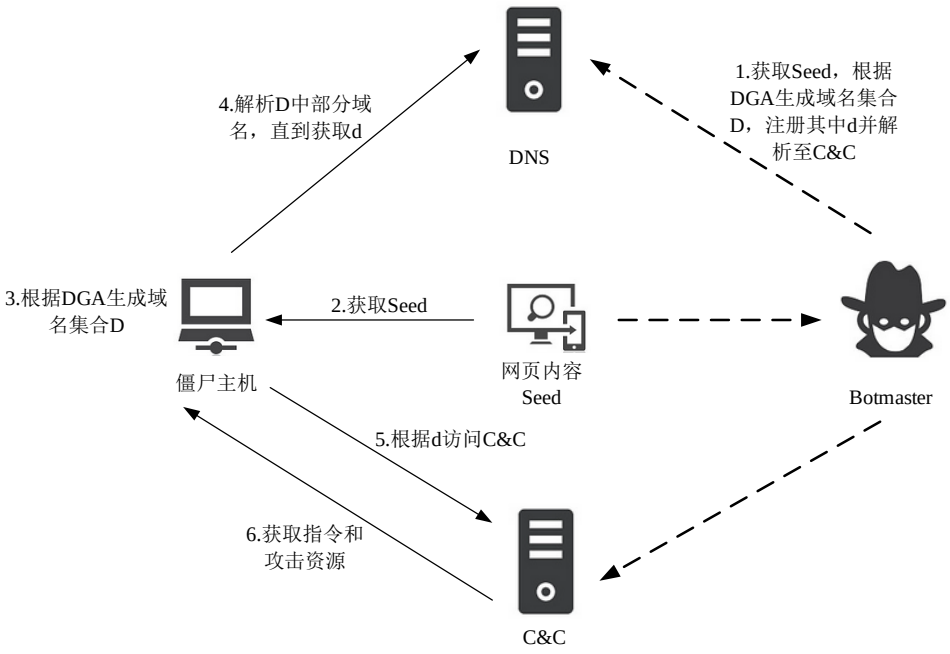


图 2. 6 Domain-Flux 命令控制流程

Domain-Flux 的核心机制是，僵尸主控机和僵尸主机共享一套 DGA 算法，二者使用相同的种子信息作为输入，生成同样的域名集合 D，僵尸主控机选择集合 D 中的部分域名 $\{d \mid d \in D\}$ 进行注册。僵尸主机在获取种子得到域名集合 D 后，每次尝试 D 中的一部分域名进行遍历，直至域名解析成功。

Domain-Flux 能够很好地抵御黑名单屏蔽等阻断措施，赋予了僵尸网络回复能力。然而，大量的 NXDomian 报文很容易暴露僵尸主机意图，在网络流量检测中，通常表现为一段时间内 DNS 访问流量暴增后又迅速下降的特征，且由 DGA 算法生成的域

名通常拥有其字符特征，使用机器学习的相关算法能够实现一定程度的识别。

2.3.5 P2P 协议

P2P 型僵尸网络是基于 P2P 协议构建命令控制信道的僵尸网络。P2P 的核心思想是对等网络，即在僵尸网络中，不存在明显意义的僵尸主机与 C&C 主机划分，僵尸网络中的所有节点既可以作为僵尸主机，也可以作为 C&C 主机，这种机制使得 P2P 僵尸网络相对于其他类型更加健壮和易扩展[37]。P2P 型僵尸网络拓扑如图 2.7 所示：

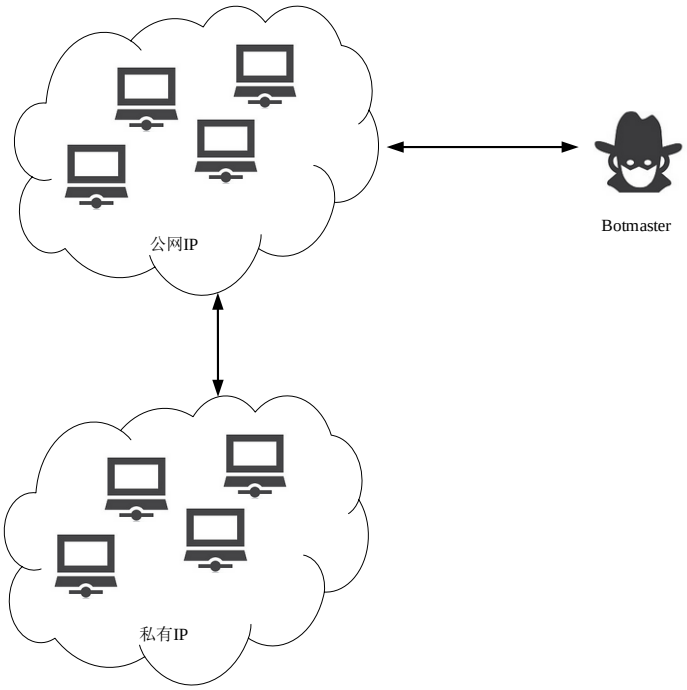


图 2. 7 P2P 型僵尸网络拓扑

P2P 型僵尸网络中拥有公网 IP 的任意节点既可以作为 C&C 主机，解决了僵尸网络的单一节点失效问题。如图 2.7 所示为典型的半分布式 P2P 型僵尸网络，根据节点功能不同，将僵尸网络节点分为两类：一类是拥有公网 IP，能够由外网发起访问的节点，称之为 Servant，这类节点在僵尸网络中既可以作为僵尸主机，也可以作为 C&C 主机，另一类是 IP 地址动态分配，IP 私有，无法从公网进行访问的节点，如校园网节点，称之为 Client。P2P 型僵尸网络中每个节点都保存有一张多个 Servant 节点的链接（Peer-List）表。当僵尸主机被迫脱离僵尸网络时，会根据链接表重新上线。

P2P 型僵尸网络相比用户 IRC 和 HTTP 型，更加健壮，但仍存在一些缺陷。首先，P2P 型僵尸网络的认证机制不完善，僵尸主控机与 C&C 的连接比较松散，没有统一的中心认证机制，因此，很容易被安全人员通过欺骗获取链接表，从而对僵尸网络进行破坏。其次，P2P 型僵尸网络的初始化极其脆弱，在 bootstrap 过程中，需要将

链接表以硬编码的方式写入僵尸网络，一旦链接表失效，则僵尸主机就会被孤立。P2P型僵尸网络比较复杂，受限于篇幅，文章只做简短的介绍。

2.4 僵尸网络对抗关键问题

结合上述僵尸网络行为及协议分析，纵观僵尸网络与安全机制对抗的历史过程，总结以往僵尸网络对抗经验，得出僵尸网络对抗的关键问题包括：

（1）单一节点失效问题，在对抗过程中，僵尸网络尽可能避免集中管控的方式，僵尸网络通过改变其拓扑结构保证整个网络的健壮性。如 P2P 型僵尸网络使用轮训机制来保证 C&C 主机的变化，即使阻断了一个 C&C 连接，还可以通过链接表中的其他 C&C 进行恢复。

（2）IP 和域名识别问题，能不能有效阻断僵尸网络的连接取决于是否能准确的识别出 C&C 的 IP 及域名。僵尸网络通过 Fast-Flux 和 Domain-Flux 技术实现 IP 和域名的隐藏，随之而来的是 DNS 解析日志的异常，安全人员通过分析异常的 DNS 流量能够识别出僵尸网络行为，却很难有效阻断僵尸网络连接。

（3）僵尸主控机隐藏问题，IRC 型僵尸网络要求僵尸主控机与僵尸主机同时接入 C&C，极大的增加了僵尸主控机暴露的风险，HTTP 型僵尸网络实现了发布指令与执行指令在时间和空间上的分离，僵尸主控机发布指令后即可断开与 C&C 的连接，等待僵尸主机定时访问获取指令，能够更好的隐藏僵尸主控机。

（4）安全机制对抗问题，IDS、防火墙、杀毒软件等安全策略能够有效地阻断僵尸主机与 C&C 的通信，僵尸网络通过使用内容加密、变换请求方式等策略进行绕过。如一般安全策略不会阻断 HTTP 流量，僵尸网络将交互信息隐藏在 HTTP 流量中，能够穿透 IDS 和防火墙。同样的还包括 DNS 隧道技术。

（5）调试分析对抗，通过样本解析和调试，安全人员能够很快分析出僵尸程序的控制协议，攻击方式，C&C 地址等信息，僵尸程序为了保护这类关键信息不被轻易获取，使用了进程标志、异常触发等各种策略防止安全人员进行样本分析。

2.5 本章小结

本章首先介绍了僵尸网络的基本概念，首先描述了僵尸网络的定义，对僵尸网络涉及的相关知识进行综述，将僵尸网络与其他恶意病毒进行了区分。其次，着重论述了僵尸网络传播、寻址、潜伏、交互和攻击等行为，重点分析了僵尸网络命令控制协议，最后，总结了僵尸网络对抗中的关键问题。

第三章　僵尸网络行为模型

僵尸网络在运行过程中，僵尸主机流量及日志会呈现出其特有的字符特征及统计规律，文章中将这些特征抽象为行为标签并作为识别和分析僵尸网络的依据。通过对这些行为标签进行统计和关联分析，实现僵尸网络画像。本章首先简述僵尸网络行为画像的概念，然后围绕行为标签，构造僵尸网络行为标签模型。最后，结合僵尸网络生命周期中不同状态的转换时序关系，提出基于 CRF 的僵尸网络识别模型。

3.1 僵尸网络画像

在安全领域，各类研究人员由于其所处的安全场景不同，对僵尸网络研究的安全诉求亦不尽相同，安全公司侧重于如何有效对僵尸网络进行防护和阻断，国家机构则侧重于对僵尸网络规模及破坏力进行检测和评估，军事领域则更侧重于在非合作环境中对僵尸主机进行溯源和取证。由于上述安全场景不同，导致了对僵尸网络的研究较为片面，文章针对僵尸网络本体，通过其行为特征研究僵尸网络画像技术，以更全面地描述僵尸网络的多维特征。

传统的僵尸网络检测机制使用数据融合或聚类等方法提取僵尸网络特征值，与僵尸网络特征库的固定字段进行规则匹配，通过计算其相似度识别僵尸主机。然而，随着僵尸网络在传播机制，交互协议，攻击类型等方面的复杂度不断提高，尤其是 APT 等高级网络攻击日益成为主流，在当前严峻的网络安全形势下对防御机制提出了更高的要求，不仅需要识别僵尸主机，还需要对僵尸主机受害情况进行评估，提供有效的阻断策略，甚至实现对僵尸主控机和 C&C 主机的溯源。传统的规则匹配得到的结果不能在宏观上对整个僵尸网络行为进行描述，导致其防御策略较为片面。因此，文章提出了基于行为分析的僵尸网络画像技术，对整个僵尸主机在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度进行描述，将僵尸主机在僵尸病毒生命周期内各个时期的散乱行为标签进行结构化整理，在多个维度中对攻击行为进行描述，最终形成僵尸网络画像。

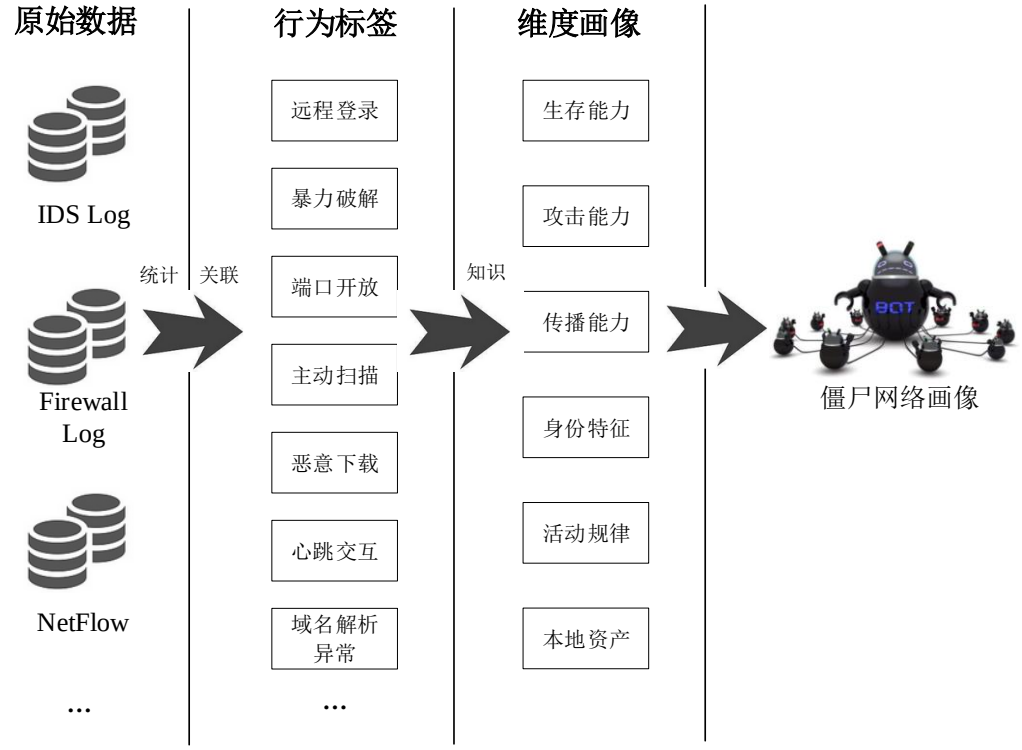


图 3. 1 僵尸网络多维画像

如图 3.1 所示为僵尸网络多维画像示意图，现阶段，随着僵尸病毒隐藏能力的不断提升，僵尸网络生存周期大幅增长，僵尸网络行为在时间和空间上的跨度不断增大，单纯从实时流量中提取的僵尸主机行为特征并不能对僵尸主机行为进行宏观描述，因此，对原始数据采集时，不光从当前网络状态和流量中采集信息，同时也从入侵检测系统日志、防火墙日志等相关历史记录中进行提取，原始数据的来源在时间跨度上力求最大。僵尸网络的行为标签不仅包含类似端口状态、远程登录登等与时间点相关的僵尸网络行为标签，还包括更多与时间段相关的行为标签，如 Mirai 的主动扫描行为，会使得一段时间内的 SYN 流量突然增加，然后又瞬间恢复；DGA 域名解析在一段时间内会存在大量的 NXDomain 信息，因此，应当为所有的行为标签均设置时间窗口和时间戳。当网络中出现类似僵尸网络行为的数据时，根据定义的标签格式，为每个行为打上相应的标签，以备后期处理。在维度画像中，通过建模分析的方法，将相互关联的标签做结构化描述，对整个僵尸网络的某个维度进行分析和画像，每一组结构化标签都对僵尸网络的一个维度进行描述，最终实现对僵尸网络在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的僵尸网络画像。

研究基于行为标签的僵尸网络画像技术，涉及特征工程，威胁情报，用户画像，知识工程等多个领域，需要网络流量采集，元数据提取，行为标签识别，关联画像等多个模块协同工作，共同构建基于行为标签的僵尸网络画像系统。其系统架构如图 3.2 所示：

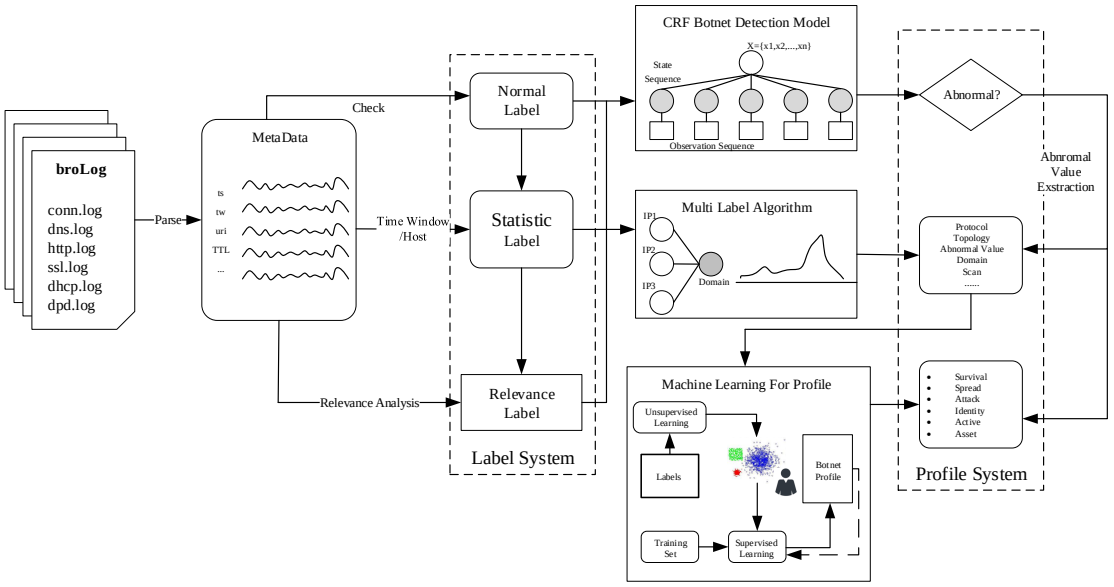


图 3.2 僵尸网络画像系统架构

如图 3.2 所示为基于行为僵尸网络画像的系统架构。主要包括了标签系统和画像系统两部分。其余的各个模块及算法均作为辅助工作出现。

（1）标签系统。行为标签是对僵尸网络行为的抽象表达，是构建整个僵尸网络画像系统的基础。文章对僵尸网络的行为标签进行了分层定义，其中 Normal Label 为基础标签，通常只需要对元数据字段的特征值进行对比即可得出。Statistic Label 为统计型标签，由于僵尸网络的大部分特征均与时间段或主机连接相关，确定了时间窗口和主机之后，能够得到大量的统计型标签。Relevance Label 为关联型标签，通过将一段时间内的多种行为特征相组合，结合关联分析等机器学习相关算法，得到僵尸网络的行为标签。

（2）画像系统。是对僵尸网络本体各个维度的属性进行充分的描述和评估结果。僵尸网络画像包括生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度。画像系统首先通过对已经标记的僵尸网络行为标签及画像结果进行学习，提取其中规则与关联关系，训练画像模型。在对僵尸主机进行画像时，通过计算行为标签相似度，推荐适配度较高的画像模型，结合实时网络环境，对僵尸主机进行画像。

图 3.2 同时描述了基于行为标签的僵尸网络画像流程。首先，对僵尸主机的原始流量进行采集，使用入侵检测系统等工具生成网络状态日志，包括连接状态日志，DNS 日志，文件传输日志，HTTP 日志等，这些日志包含了提取僵尸网络行为标签的所有元数据。然后，根据标签生成的相关算法，设置时间窗口和连接主机，分别得出在各个时间段的基础标签，统计标签和关联标签，为了保证画像系统的开销尽可能不影响系统的正常业务，将标签进行层次划分，基于元数据特征的行为标签最容易获取，在捕捉到此类行为时，才对日志文件进行关联分析等耗时操作。获取行为标签后，使用

基于 CRF 的僵尸网络行为模型对僵尸主机进行识别，判断其是否属于僵尸主机。最后，结合僵尸主机的属性和规则，根据先验知识对僵尸主机在生存能力、攻击能力、传播能力、身份特征、活动规律和本地资产六个维度的画像。

3.2 基于 CRF 的僵尸网络识别模型

在对僵尸网络进行多维画像前，首先要对僵尸主机进行识别。传统的检测方法包括基于通信内容的检测机制，基于日志挖掘的检测机制。基于通信的检测机制通过识别通信内容中特定的僵尸网络特征值或通信流量时序特征识别僵尸网络通信流量，这种方法需要在防火墙或 IDS 上预先配置好相应的匹配规则，因此只能识别出已有的僵尸网络通信，检测准确度较高。基于日志挖掘的检测机制通过识别日志文件中的异常值，并结合异常值之间的关联关系识别僵尸网络行为。文章提出基于 CRF 的僵尸网络识别模型，该模型基于僵尸网络生存模型，将僵尸网络的生存周期划分为感染、潜伏、维护、传播、攻击、销毁六个阶段，各个阶段之间存在时序关系，通过识别僵尸网络行为特征，将其映射到所属阶段的生命周期模型中，结合生命周期各阶段的时序特征，达到识别僵尸主机的目的，该种方法能够有效地发现未知类型僵尸网络和已知僵尸网络变种。

3.2.1 僵尸网络生存模型

通过对现阶段若干种僵尸网络进行分析，提出僵尸网络生存模型，将僵尸网络根据生存周期分为感染，维护、潜伏、传播、攻击、销毁六个状态，其状态转换关系如图 3.3 所示：

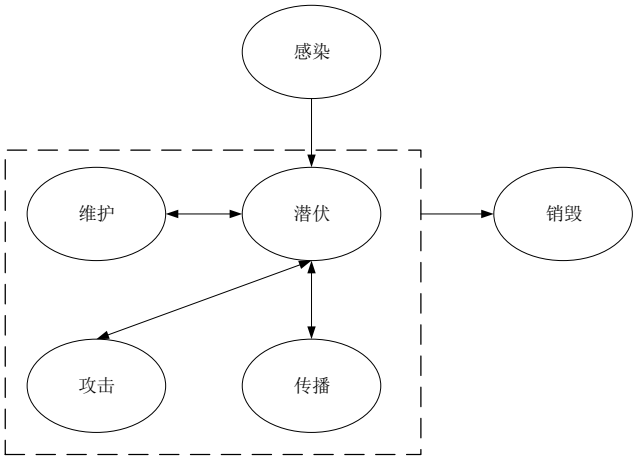


图 3. 3 僵尸网络状态图

如图 3.3 所示，感染态与销毁态为不可逆状态，维护态、攻击态和传播态均可以通过潜伏状态进行转换，在转换过程中，均会出现寻址行为。判断僵尸网络所处状态，并结合状态转换时序关系，结合其他状态行为标识，能够达到识别僵尸网络的目的，

僵尸网络各个状态下的行为特征如表 3.1 所示：

表 3. 1 僵尸网络状态描述

僵尸网络状态	状态描述	常见可疑行为	常见特征
感染状态	用户通过多种渠道感染僵尸病毒，被远程登录并主动下载攻击载荷，与 C&C 进行特殊认证，成为僵尸主机	1） 端口异常打开 2） 远程登录 3） 暴力破解 4） DNS 流量异常 5） 创建目录 6） 异常下载	1） 登录失败日志异常 2） 端口日志异常 3） 域名解析频繁，DNS 流量异常 4） 注册表修改 5） FTP、wget 下载 6） 端口强制打开或关闭 7） 域名请求结果不一致
潜伏状态	僵尸主机周期性与 C&C 进行心跳确认，保证在线	1） 周期性通信 2） P2P 通信	1） PING/PONG 通信 2） 周期性特殊内容通信 3） 周期性解析相同域名
传播状态	僵尸病毒利用漏洞、社会工程、介质摆渡等方法进行传播	1） 全网扫描 2） 漏洞利用 3） 远程登录 4） 异常上传	1） 扫描流量 2） IP 伪装远程登录 3） POST 上传
维护状态	僵尸网络获取载荷更新，隔离后重新寻址备用 C&C	1） 获取链接表 2） 异常下载 3） DNS 流量异常	1） 域名解析频繁，DNS 流量异常 2） P2P 流量异常 3） 异常下载
攻击状态	僵尸网络根据僵尸主控机的指令进行邮件发送，DDoS 攻击，敏感信息窃取，资源滥用等攻击行为	1） DDoS 攻击 2） 用户操作记录 3） 恶意点击欺诈 4） 比特币挖矿 5） 恶意邮件发送	1） TCP、UDP、ICMP 等流量异常 2） 建立代理邮件服务 3） 性能下降 4） 后台进程实例异常 5） DNS 流量异常 6） 广告点击异常 7） 可用资源减少
销毁状态	僵尸病毒在主机上被移除，一般受到杀毒软件查杀或被隔离和调试时主动销毁	1） 计时探针 2） 逻辑识别 3） 蜜罐识别	1） 主动删除

3.2.2 基于 CRF 的僵尸网络行为模型

随着僵尸网络复杂程度的不断提高，传统的使用基于规则匹配的识别方法由于其片面性已经不能满足当前僵尸网络识别的要求，由于僵尸网络行为在时间和空间上呈现出诸多的序列化特性，文章基于僵尸网络的生存模型，文章构建了基于 CRF（Condition Random Field 条件随机场）的僵尸网络行为模型，用以解决僵尸网络的识别问题。

将僵尸网络生存模型抽象为如图 3.4 所示的图模型，节点表示僵尸网络生命周期所处的状态，边表示状态之间的转换概率。其次，在 CRF 模型中，通过对已知僵尸网络行为分析，获取到其状态序列，针对所有的状态序列，使用特征函数训练出在已知状态序列的前提下，图中各个节点之间的联合概率分布，并将每个序列的概率分布进行线性表示和归一化，求得各个状态序列的分值。最后，在识别过程中，针对状态序列，使用相同的特征函数求得该序列的分值，当分值大于特定阈值时，则该序列被识别为僵尸网络序列。

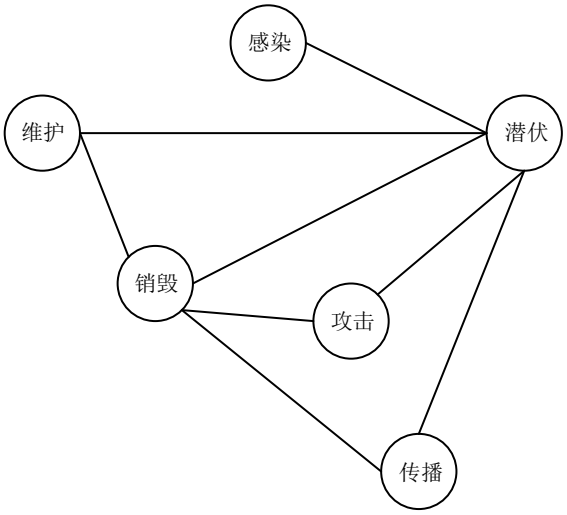


图 3.4 僵尸网络状态图模型

CRF 是一种判别式无向图，对于解决序列识别和预测问题有良好的效果。在僵尸网络的识别过程中，由于僵尸主机行为的多样性，通过识别僵尸网络的单一行为特征无法达到准确识别僵尸网络的目的。在文章中使用 CRF 模型，通过识别僵尸网络行为，判断僵尸网络状态，针对僵尸主机状态转换序列，识别僵尸主机，能够有效地减少僵尸主机误报。其模型如图 3.5 所示：

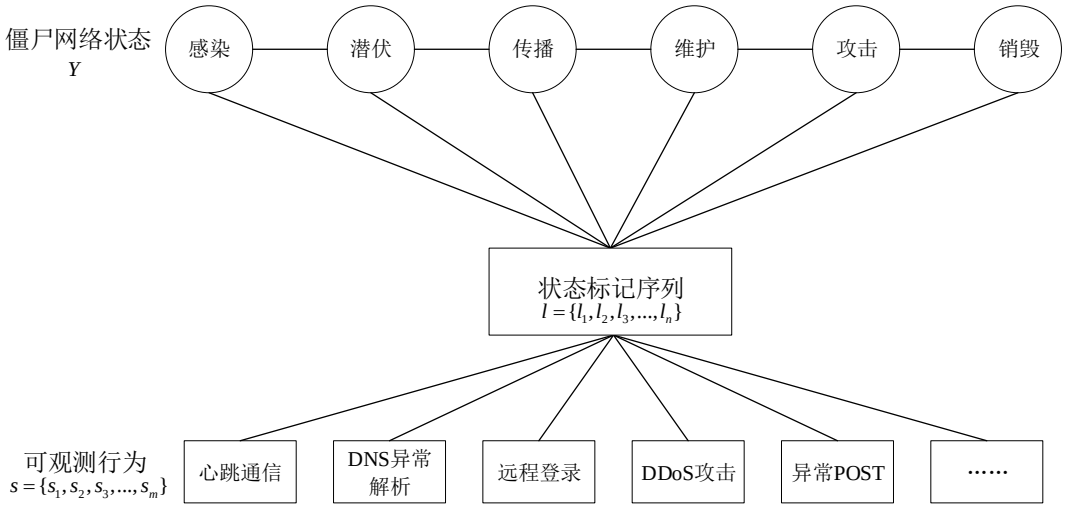


图 3. 5 基于 CRF 的僵尸网络行为标签模型

如图 3.5 所示。基于 CRF 的僵尸网络识别过程如下：

（1）获取观测行为序列。通过在网络流量分析其中配置策略，识别出僵尸网络的特征行为，如心跳通信，远程登陆，DDoS 攻击等，将其表示为$S=\{s_1,s_2,s_3,.....,s_n\}$。

（2）获取状态标记序列。通过对已知僵尸网络行为进行分析，建立观测行为与僵尸网络生命周期状态之间的对应关系。对于观测序列，结合已有的僵尸网络分析结果，将观测序列映射到僵尸网络的感染、潜伏、维护、攻击和销毁六个状态中，得到状态标记序列$L=\{l_1,l_2,l_3,.....,l_m\}$。

（3）计算已知僵尸网络状态序列概率联合分布。针对已知的僵尸网络状态序列，计算在已知序列条件下每个节点的分值：

$$score(l\mid s)=\sum_{j=1}^{m}\sum_{i=1}^{n}\lambda_j f_j(s,i,l_i,l_{i-1}) \qquad (3\text{-}1)$$

其中，λ表示该序列中，状态对应的权值，f为特征函数，使用公式（3-1）计算出该序列中所有节点的分值，并对所有的分值进行线性表示，最后进行归一化，将序列分值映射在（0,1）范围内。

（4）对已知的所有僵尸网络状态序列重复上述的打分过程，对于状态序列中的节点，根据其对判定序列是否属于僵尸网络行为序列的贡献，设置不同的权值，并结合大量已知样本训练出每个状态转移序列中节点的权值。维护状态转移概率矩阵。当输入新的状态序列时，根据已经得到的节点权值，使用相同的特征函数，依照公式（3-2）计算输入序列的分值。

$$p(l \mid s) = \frac{\exp[score(l \mid s)]}{\sum_{l'} \exp[score(l' \mid s)]} \quad (3\text{-}2)$$

针对已经检测到的状态序列，使用相同的特征函数进行计算，当最终得到的概率高于特定阈值时，该序列被标记为僵尸网络。

在对僵尸网络进行识别时，引入了 CRF 模型，其优势在于在进行恶意行为识别时，同时捕捉到僵尸网络相邻状态行为特征的概率较小，HMM（Hidden Markov Model 隐马尔科夫模型）模型[38]虽然也能做到序列识别，当观测恶意行为在僵尸网络状态上距离较远时，HMM 无法进行识别。同时，CRF 对所有的状态转换关系均进行概率计算，CRF 能够统计整个标注序列的全局概率，其打分机制更具有全局特性。

3.3 本章小结

本章首先描述僵尸网络画像的基本概念和机制，其核心是构造基于僵尸网络行为标签的画像模型，然后，探讨了僵尸网络的生存周期模型，最后，以生存周期为基础，结合僵尸网络状态转换的时序关系，提出了基于 CRF 的僵尸网络识别模型。

第四章 僵尸网络行为标签研究

研究基于行为分析的僵尸网络画像，其核心是僵尸网络行为标签体系，僵尸网络行为标签需要充分考虑僵尸网络行为在时间和空间的关联特性、标签定义形式及内容、标签时间窗口等因素。本章首先以 Mirai 这一典型僵尸病毒为例，提取其行为特征，结合多种病毒分析结果，归纳出描述僵尸病毒行为的元数据，然后，根据僵尸网络行为分析的要求，设计行为标签结构，最后，结合僵尸网络与安全机制对抗关键问题，针对实际情况，使用不同方法识别僵尸网络行为标签。

4.1Mirai 僵尸网络行为

Mirai 是一种主要针对物联网设备的僵尸病毒。随着物联网行业的发展，越来越多的 IoT 设备被接入互联网，Mirai 利用 IoT 设备默认弱口令这一漏洞进行传播，使用暴力破解的方式远程登陆 IoT 设备，并对设备进行控制，该病毒融合多种扫描传播及恶意攻击手段。Mirai 因其在 2016 年导致的美国和德国大规模断网事件而备受关注，也引发了安全研究人员和社会公众对于 IoT 设备安全的思考。Mirai 作为近年来最具影响力的僵尸病毒，最能代表现阶段僵尸病毒的行为特征以及演变规律。因此文章选取具有代表性的 Mirai 僵尸病毒进行行为分析[39]。

4.1.1 行为捕捉

对 Mirai 僵尸病毒行为进行分析，提取 Mirai 僵尸网络的细粒度行为标签，首先需要僵尸病毒在僵尸主机上的行为原始数据作为分析素材[40]，如图 4.1 所示为 Mirai 僵尸网络行为捕捉的实验环境，根据对其源码的分析，搭建了 C&C，Loader 和 DNS 三个服务，DNS 负责解析 C&C 和 Loader 服务的域名，使用网络摄像头作为待感染设备，实验过程中对 IoT 设备流量进行旁路收集，将感染设备流量作为 Mirai 僵尸网络行为分析的原始数据，进一步提取其行为标签。

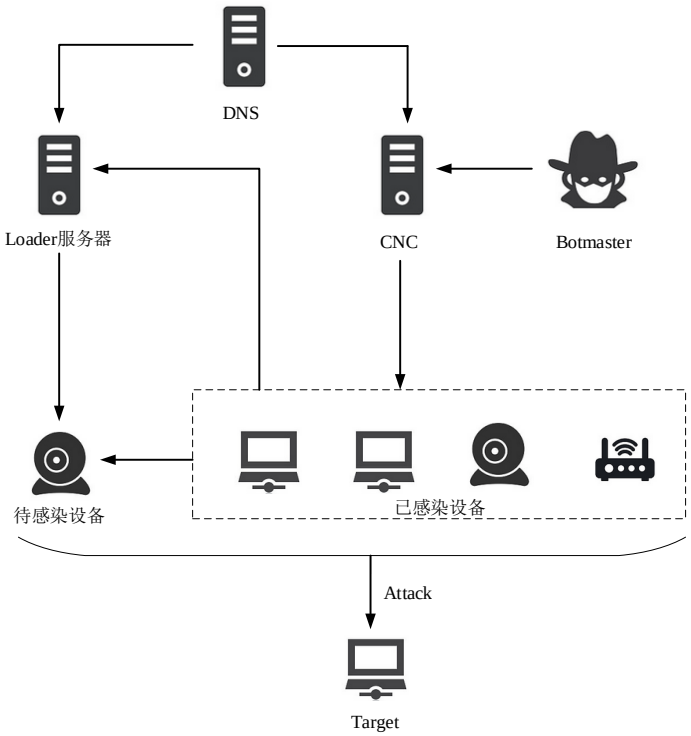


图 4. 1 Mirai 僵尸病毒行为提取实验环境

如图 4.1 所示为 Mirai 僵尸病毒实验环境，为了保证僵尸主机数据的完备，将 Mirai 源码在 debug 模式下默认关闭的 scanner 模块激活，开启 Mirai 僵尸病毒的扫描传播功能，并关闭受攻击主机的 23 端口，保证其不被感染。由于 Mirai 病毒具有极强的传播能力和破坏能力，将整个实验环境部署在物理隔离的同一个网段中，避免造成事故。实验中受感染的设备使用国产某品牌网络摄像头，该设备在针对 DYN 的攻击中提供了约 30 万个僵尸主机。

Mirai 僵尸主机使用了特殊的感染机制，可以在僵尸主控机极少参与的情况下快速扩充 Mirai 僵尸网络规模，僵尸主机对网段中的所有网络设备进行 SYN 扫描，尝试使用 61 组弱口令进行 23 端口暴力破解登陆，若成功，则将漏洞主机的相关信息存储在 ScanList 表中，定时发送给 Loader 服务器，由 Loader 服务器对漏洞主机进行远程登陆，并操作漏洞主机下载 Mirai 病毒，成功感染后，漏洞主机会对僵尸病毒中硬编码的 C&C 主机域名进行解析，发送特殊认证消息进行僵尸主机上线操作。这种感染机制的效率极高，Mirai 作者 Anna-senpai 在 Hack Forums 公布 Mirai 源码时曾指出，该方法每秒会得到 500 个成功爆破的结果，且二次感染几率极高。基于这种传染机制，在实验环境中搭建了 DNS 服务器、Loader 服务器和 C&C 服务器，DNS 对 Loader 和 C&C 服务器进行解析。Mirai 除了提供一般的 TCP，UDP 攻击，还包括了效率更高的 GRE 攻击。

4.1.2 行为分析

为保证 Mirai 僵尸网络行为数据的完备性和真实性，在实验环境中复现了 Mirai 僵尸病毒的所有行为，截取对应的流量数据，结合实验中采集的行为数据和 Mirai 僵尸病毒源码进行 Mirai 僵尸主机行为分析，获取 Mirai 僵尸主机在生命周期不同状态下的主要行为及获取方式，如表 4.1 所示为 Mirai 僵尸主机行为分析结果。

表 4. 1Mirai 僵尸主机行为分析

Mirai 行为		特征描述	获取方式	区分度
感染	暴力破解	使用 61 个弱口令组合进行暴力破解，表现为一段时间出现大量登陆失败日志	Failed_login	高
	Telnet 登录	远程登陆 23 端口，使用 Telnet 服务	Port	中
	关闭 23 端口	将 23 端口关闭，防止其他进程使用	Port	中
	打开 48101 端口	防止多实例运行，打开 48101 端口作为互斥机制	Port	高
	喂狗	防止 IoT 设备重启	-	低
	排他	杀死其他僵尸病毒，独占资源	Kill	低
	信息上传	上传设备信息	POST	低
	下载病毒	下载同设备匹配的僵尸病毒	GET	低
	上线	使用特定字符认证双方	TCP	高
传播	SYN 扫描	扫描网段中的其他设备	SYN	高
	上传 ScanList	上传包含漏洞设备信息	POST	中
寻址	寻址 Loader	解析硬编码，DNS 请求	DNS	低
	寻址 C&C	解析硬编码，DNS 请求	DNS	低
潜伏	心跳连接	维持主机与 C&C 连接	TCP	高
交互	信息上传	上传设备相关信息	POST	中
	上传 ScanList	上传包含漏洞设备信息	POST	中
	下载病毒	下载同设备匹配的僵尸病毒	FTP，wget	中
	攻击指令	接收来自 C&C 的攻击指令	TCP	低
攻击	ATK_VEC_UDP	UDP 攻击，可见大流量访问	UDP	中
	ATK_VEC_VSE	大流量访问	TCP	中
	ATK_VEC_DNS	大量 DNS 请求无应答	DNS	中