# 缓冲区溢出

北京邮电大学

网络空间安全学院



姓 名 林于翔

学 号 **2020211919**

班 级 2020211806

任课教师 徐国胜

2022 年 10 月 24 日

## 摘要

如今，信息技术高速发展，各式各样的软件如雨后春笋一般出现。然而，由于软件自身存在的漏洞而被攻击的事件也屡见不鲜，缓冲区溢出漏洞便是其中的代表之一。伴随信息安全技术的不断发展，缓冲区溢出漏洞的检测技术也在不断发展。该文将首先介绍有关缓冲区及缓冲区溢出漏洞，再阐述近几年关于缓冲区溢出漏洞检测的新技术，最后是对关于缓冲区未来可能发展的探讨。

## 关键词

缓冲区溢出漏洞；缓冲区；漏洞检测

## 引言

缓冲区溢出是一种非常普遍、非常危险的漏洞，在各种操作系统、应用软件中广泛存在。利用缓冲区溢出攻击可以导致程序运行失败、系统宕机等严重后果。更严重的是，可以用它来执行非授权指令，甚至取得系统权限，从而进行非法操作。

攻击者通过往程序的缓冲区写超出其长度的内容，造成缓冲区溢出，从而破坏程序的堆栈，造成程序崩溃或程序转而执行其他指令，已达到攻击的目的。[1]造成缓冲区溢出的原因是程序中没有仔细检查用户输入的参数。

## 缓冲区溢出的研究

早在20多年前缓冲区溢出漏洞就被用来配合蠕虫病毒进行扩散，只是当时人们并没有对缓冲区溢出问题加以重视。直到1996年AlephOne详细描述了如何在linux中利用栈溢出漏洞[2]，这是缓冲区溢出漏洞技术在真正被人们所认识。

在1998年Lincoln实验室用来评估入侵检测的的5种远程攻击中，有2种是缓冲区溢出。而在1998年CERT的13份建议中，有9份是是与缓冲区溢出有关的，在1999年，至少有半数的建议是和缓冲区溢出有关的。在Bugtraq的调查中，有2/3的被调查者认为缓冲区溢出漏洞是一个很严重的安全问题。

在21世纪的今天，缓冲区溢出漏洞仍然是影响当前网络安全的最主要问题之一[3-6]。

根据微软公司2010年发布的106项危险漏洞，有2/3以上的危险漏洞与缓冲区溢出漏洞有关。在目前已经确认的软件漏洞中，缓冲区溢出漏洞约占到20%。[7]在ICAT 2003-9-12号公布的10大漏洞 排名中居于榜首的仍然是缓冲区溢出漏洞；从 CERT公布的截止2005年为止的统计数据中可以看出，缓冲区溢出攻击的问题正在扩大化。

根据CNCERT报告，在2011年国家信息安全漏洞共享平台CNVD共收集整理并公开发布信息安全漏洞5547个。在所有漏洞中，涉及各种应用程序的比例为62.6%，涉及各类网站系统的漏洞比例为22.7%，而涉及各种操作系统的漏洞则占了8.8%。在这些安全漏洞中明确是缓冲区溢出漏洞的数量就有近500[8]个。因此本文针对近几年常见的缓冲区溢出漏洞技术进行整理。

## 常见的缓冲区溢出检测技术

静态检测技术

静态检测技术无需运行程序，消耗资源少，检测对象一般是程序源代码，多数情况下可以准确地定位到程序所在位置，便于编程人员修改，从而彻底将检测出的漏洞消除。

Peng Luo[9]等人研究发现有三分之一的缓冲区溢出漏洞是因为循环，因此他们使用CLI实现Dbloop来定位可以引起缓冲区溢出的循环。其关键步骤是获取缓冲区大小，然后检查目标循环的数据移动是否超出了缓冲区的范围。

Lili Xu[10]等人重点研究整数溢出导致的缓冲区溢出（IO2BO）漏洞。他们提出一种新的静态分析框架（LAID），首先利用过程间数据流分析和污点分析来准确是被潜在的漏洞，然后使用轻量级方法进一步滤除误报。实验结果表明，他们提出的新的框架可以有效地检测所有一直的IO2BO漏洞。

周艳艳[11]积极借鉴机器学习的思想和方法，并采用随机森林算法构建多棵决策树对样本进行训练并预测。

动态测试技术

动态测试技术是指通常无限大的执行域中恰当地选取一组有限的测试用例来运行程序，从而检测程序的实际运行结果是否符合预期结果的分析手段[12]。动态测试技术的核心在于如何生成覆盖率较高的测试用例，或者是命中要害、触发缓冲区溢出的测试用例。

Rui Zhang[13]等人对卫星软件中堆栈溢出的普遍问题，对RTEMS操作系统进行了改进，以支持实时堆栈使用深度和溢出检测。改进的RTEMS通过上下文切换在线程之间共享访问保护机制。陷阱处理程序旨在接管写保护错误陷阱，计算堆栈使用深度并实时监视堆栈溢出。

## 未来研究热点展望

本文认为未来针对缓冲区溢出检测技术的主要热点应在于以下三个方向：

结合机器学习的思想[14]。机器学习的主要思想在于通过不断地训练，提取出训练集的特征，并达到精确预测的目的。而缓冲区溢出漏洞有规律可循，可以将代码根据数组的定义、指针的声明等分解为若干片段，利用机器学习进行训练，减少测试中的误报和漏报问题。

将高级语言转化为底层语言如汇编或二进制代码[15]。缓冲区溢出漏洞常常发生在C/C++，对于汇编或二进制代码检测技术却少之又少。如果能有针对汇编二进制代码的检测技术，将极大减少误报率。

综合运用各项技术。将动态检测技术和静态检测技术相结合，发挥各自的优势，将更有利于防范缓冲区溢出漏洞。

## 参考文献

[1] 罗斌.缓冲区溢出漏洞分析及防范[J].电脑知识与技术,2018, 14(33):44-45.

[2] AlephOne.Smashing The Stack For Fun And Profit [J]. Phrack,1996,7(49)

[3] 邓爽缓冲区溢出攻击分析及防范策略研究［Ｄ］．济南：山东大学，2009

[4] 李毅超，刘丹，韩宏，等缓冲区溢出漏洞研究与进展［Ｊ］．计算机科学，2008，35(1)：87-89,125

[5] 林志强，夏耐，茅兵，等缓冲区溢出研究综述［Ｊ］.计算机科学，2004，31(9):110-113,160

[6] 王业君，倪惜珍，文伟平，等．缓冲区溢出攻击原理与防范的研究［Ｊ］．计算机应用研究,2005,22(10)：101-104

[7] Forst JC.Osipov V,Bballa N,et al.Buffer overflow attacks :detect, exploit, prevent[M]. Rockland: Syngress Press,2005.

[8]2011年我国互联网网络安全态势综述[EB/OL].http://www.cert.org.cn/UserFiles/File/201203192011annualreport.pdf

[9] Peng Luo, Deqing Zou, Yajuan Du, Hai Jin, Changming Liu, Jinan Shen. Static detection of real-world buffer overflow in⁃ duced by loop. Computers & Security, 2020.

[10] Mingjie Xu, Shengnan Li, Lili Xu, Feng Li, Wei Huo, Jing Ma, Xinhua Li, Qingjia Huang. A Light-Weight and Accurate Method of Static Integer-Overflow-to-Buffer-Overflow Vulner⁃ ability Detection[J]. Springer Nature Switzerland AG. 2019.

[11] 周艳艳. 基于随机森林算法的缓冲区溢出攻击检测研究 [D]. 哈尔滨: 哈尔滨理工大学, 2018.

[12] Patton P. Software Testing. 2nd ed., SAMS, 2005.

[13] Rui Zhang, Yan Du, Tao Zhang, Qi Qiu, Liang Mao, Jiaxiang Niu. An Improved RTEMS Supporting Real-Time Detection of Stack Overflow. Springer Nature Switzerland AG. 2019.

[14]贾疏桐,桂灿,苏星宇.缓冲区溢出漏洞分析及检测技术进展[J].电脑知识与技术,2020,16(13):57-59.DOI:10.14004/j.cnki.ckt.2020.1615.

[15] 邵思豪,高庆,马森,等.缓冲区溢出漏洞分析技术研究进展[J]. 软件学报,2018,29(5):1179-1198.