

第一章 Web安全概述

世界观：互联网本来是安全的，自从有了研究安全的人之后，物联网就变得不安全了

Hacker想拿到root权限用编写的exploit进行攻击

- 有的黑客：精通技术，挖掘漏洞，编写exploit
- 有的黑客：没有动手能力，脚本小子，这些人是**主要威胁**

黑客攻击的目标：

- 早期互联网：SMTP、POP3、FTP、IRC等协议的服务拥有绝大多数用户，黑客主要目标是网络、操作系统以及软件
- Web1.0时代：**服务器**动态脚本安全问题（SQL注入，XSS）
- Web2.0时代：**客户端**安全问题

白帽子与黑帽子

黑帽子：利用黑客技术搞破坏

白帽子：精通安全技术反黑客

对于黑帽子来说：只要找到一个系统弱点，就能完成入侵

对于白帽子来说：必须找到所有系统弱点，才能完成防御

“破坏永远比建设容易”

扭转局面：白帽子需要克服某种攻击方法，而非抵御单次攻击，**化被动为主动**

- 例：设计一个解决方案，在特定环境下抵御所有已知和未知的SQL Injection

“最大的漏洞是人”

安全的本质是信任的问题

使用**信任边界**划分**信任域**

- 数据从高等级信任域流向低等级信任域，无需经过安全检查
- 数据从低等级信任域流向高等级信任域，需要经过安全检查

安全方案建立在信任的基础上，必须要信任一些东西，安全方案才得以建立

极端的条件意味着**小概率以及高成本**，成本有限的前提下：

- 根据成本设计安全方案
- 可能性大的条件作为决策的主要依据

一旦决策依据的条件被打破，安全假设的前提条件就不再可靠

安全方案设计难点：把握住信任条件的度

安全三要素

- 机密性（Confidentiality）：数据内容不能泄露（加密）
- 完整性（Integrity）：数据内容不能被改写、丢失（数字签名）
- 可用性（Availability）：数据内容不能无法获得（DoS）

安全评估

资产等级划分-威胁分析-风险分析-确认解决方案

资产等级划分：确认保护对象

- 互联网的核心是由**用户数据驱动**
- 确认要保护的对象：用户数据、员工数据
- 划分信任域和信任边界
 - 简单网站信任模型：数据库最内侧、Web应用再外侧、不可信任的Internet再最外侧

威胁分析：找出所有威胁确定攻击面（Attack Surface）

- 微软提出STRIDE威胁建模方法

威胁	定义	对应的安全属性
Spoofing（伪装）	冒充他人身份	认证
Tampering（篡改）	修改数据或代码	完整性
Repudiation（抵赖）	否认做过的事情	不可抵赖性
InformationDisclosure（信息泄露）	机密信息泄露	机密性
Denial of Service（拒绝服务）	拒绝服务	可用性
Elevation of Privilege（提升权限）	未经授权获得许可	授权

风险分析：Risk = Probability * Damage Potential

- 微软提出DREAD模型

等级	高（3）	中（2）	低（1）
Damage Potential	获取完全验证权限；执行管理员操作；非法上传文件	泄露敏感信息	泄露其他信息
Reproducibility	攻击者可以随意再次攻击	攻击者可以重复攻击，但有时限制	攻击者很难重复攻击过程
Exploitability	初学者在短期内能掌握攻击方法	熟练的攻击者才能完成这次攻击	漏洞利用条件非常苛刻
Affected users	所有用户，默认配置，关键用户	部分用户，非默认配置	极少数用户，匿名用户
Discoverability	漏洞很显眼，攻击条件很容易获得	在私有区域，部分人能看到，需要深入挖掘漏洞	发现该漏洞极其困难

设计安全方案：能够有效解决问题；用户体验好；高性能；低耦合；易于扩展与升级

设计安全方案时的技巧

Secure By Default 原则含义一：白名单黑名单思想，更多的使用**白名单**，系统会更安全

- 白名单：只允许使用XXX，XXX.....
- 黑名单：不允许使用XXX，XXX.....

XSS Filter对HTML标签匹配规则，这个规则：

- 如果是黑名单：以后有新的HTML标签则不在黑名单中
- 如果是白名单：只允许特定标签可用，更加安全

Secure By Default 原则含义二：最小权限原则，授予主体必要的权限，**不过度授权**

- Linux系统使用普通账户登陆，sudo实现root权限操作

Defense In Depth 原则含义一：在不同层面、不同方面实施安全方案，避免疏漏

Defense In Depth 原则含义二：在解决根本问题的地方实施针对性的安全方案

不可预测性原则：微软ASLR技术，程序每次启动时，进程的栈基址都不相同，有一定随机性

- id=1000, id=1002, id=1003：攻击者轻易遍历并删除
- id=kzce, id=jchs, id=qwed：攻击者只能一一分析，提高了攻击门槛