北京邮电大学

网络空间安全学院



**软件详细设计报告**

**项目： KVM虚拟机镜像安全扫描工具**

**组员： 王硕、李懿飞、李朝恒**

**2021年5月11日**

目录

[1 引言 4](#_Toc58503450)

[1.1 目的 4](#_Toc58503451)

[1.2 背景及范围 4](#_Toc58503452)

[1.3 定义、术语或缩略语 4](#_Toc58503453)

[2 软件系统结构 5](#_Toc58503454)

[2.1 需求概述 5](#_Toc58503455)

[2.2 子模块划分 5](#_Toc58503456)

[2.3 子模块关系 6](#_Toc58503457)

[3 各模块详细设计 8](#_Toc58503458)

[3.1 DNS数据包捕获模块 8](#_Toc58503459)

[3.1.1 模块功能描述 8](#_Toc58503460)

[3.1.2 利用DNS隧道工具捕获恶意数据包 8](#_Toc58503461)

[3.1.3 实时检测中的DNS流量获取 11](#_Toc58503462)

[3.1.3.1 思路 11](#_Toc58503463)

[3.1.3.2 算法 11](#_Toc58503464)

[3.1.4 输入输出与性能 12](#_Toc58503465)

[3.2 PCAP文件解析模块 13](#_Toc58503466)

[3.2.1 模块功能描述 13](#_Toc58503467)

[3.2.2 PcapParser类 14](#_Toc58503468)

[3.2.3 输入 21](#_Toc58503469)

[3.2.4 输出 21](#_Toc58503470)

[3.3 DNS指纹生成模块 23](#_Toc58503471)

[3.3.1 模块功能描述 23](#_Toc58503472)

[3.3.2 DNS特征提取 23](#_Toc58503473)

[3.3.3 DNS特征解释 27](#_Toc58503474)

[3.3.4 模块输出 34](#_Toc58503475)

[3.4 机器学习特征分析模块 35](#_Toc58503476)

[3.4.1 模块功能描述 35](#_Toc58503477)

[3.4.2 输入 35](#_Toc58503478)

[3.4.3 算法 36](#_Toc58503479)

[3.4.4 特征重要度分析 42](#_Toc58503480)

[3.4.5 模型调优 42](#_Toc58503481)

[3.4.6 准确率 43](#_Toc58503482)

[3.4.6.1 混淆矩阵 43](#_Toc58503483)

[3.4.6.2 准确率 43](#_Toc58503484)

[3.4.7 模型保存 44](#_Toc58503485)

[3.5 机器学习分类模块 45](#_Toc58503486)

[3.5.1 模块功能描述 45](#_Toc58503487)

[3.5.2 算法 45](#_Toc58503488)

[3.5.3 输出 46](#_Toc58503489)

[3.6 辅助输出模块 47](#_Toc58503490)

[3.6.1 模块功能描述 47](#_Toc58503491)

[3.6.2 算法 47](#_Toc58503492)

[3.6.3 输入输出 53](#_Toc58503493)

# 1 引言

### 1.1 目的

编写该详细设计说明书是为说明该项目——KVM虚拟机镜像安全扫描工具的设计考虑，包括程序描述、输入/输出、算法和流程逻辑等，为软件编程和系统维护提供基础。本说明书的预期读者为系统设计人员、软件开发人员、软件测试人员和项目评审人员。其中系统设计人员、软件开发人员、软件测试人员为小组内部成员组成，软件评审人员为课程老师或助教。

### 1.2 背景及范围

* + 项目名称：KVM虚拟机镜像安全扫描工具
  + 项目成员：北京邮电大学网络空间安全学院“信息安全综合实践”课程开发小组
    - 王硕：2018213641
    - 李懿飞：2018213632
    - 李朝恒：2018213642
  + 系统范围：具有python3环境的CentOS7 64位系统计算机/虚拟机
  + 用户：无限制
  + 实现项目的计算机网络：校园网

### 1.3 定义、术语或缩略语

|  |  |  |
| --- | --- | --- |
| 序号 | 术语或缩写 | 解释 |
| 1 | KVM |  |
| 2 | Image（镜像） |  |
| 3 | Libguestfs |  |
| 4 | Row格式 |  |
| 5 | Qcow2格式 |  |
| 6 | NVD CVE Data Feeds |  |
| 7 | VirusTotal |  |
| 8 | dpkg包管理系统 |  |
| 9 | rpm包管理系统 |  |
| 10 | Mount（挂载） |  |
| 11 | Unmount（解挂载） |  |

# 2 软件系统结构

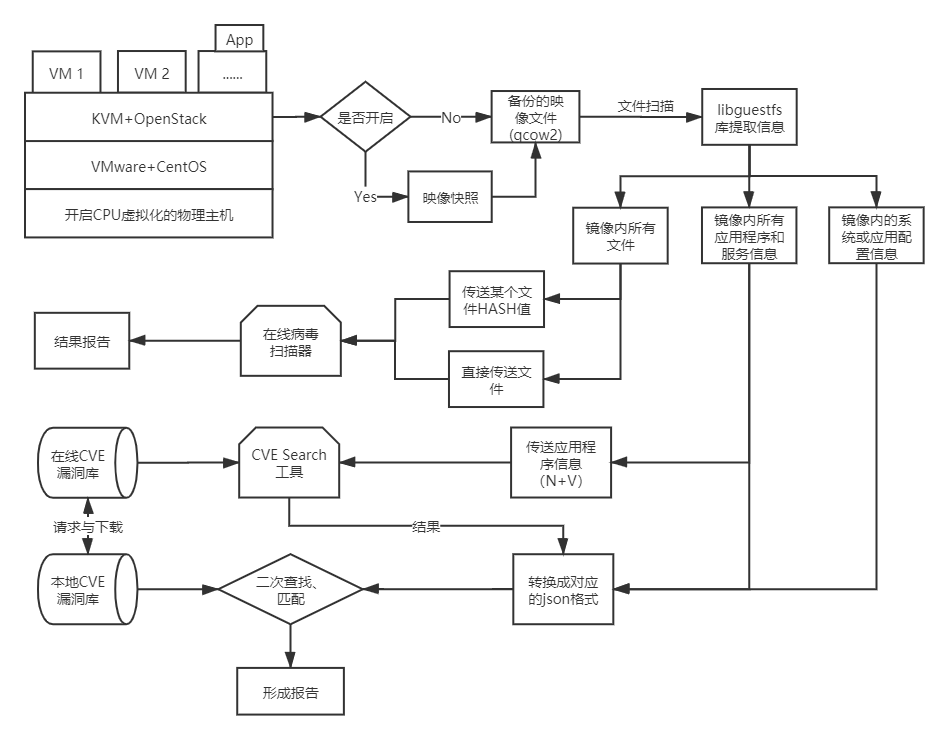
### 2.1 需求概述

开发一个虚拟化层的云安全KVM虚拟机镜像扫描工具，完善OpenStack工具在安全方面的功能，为云设备安全提供保障。

以KVM虚拟机镜像文件作为输入，通过静态解析与扫描，提取虚拟机内所有的文件、安装的应用程序和服务信息、系统和程序的各种配置信息，然后通过在线杀毒网站进行病毒扫描、通过CVE开发平台下载漏洞数据库，对比漏洞库和镜像内提取的信息，判断是否存在漏洞，最终生成包含漏洞信息和建议方案的扫描报告。

### 2.2 子模块划分

本工具的系统功能图如下所示：

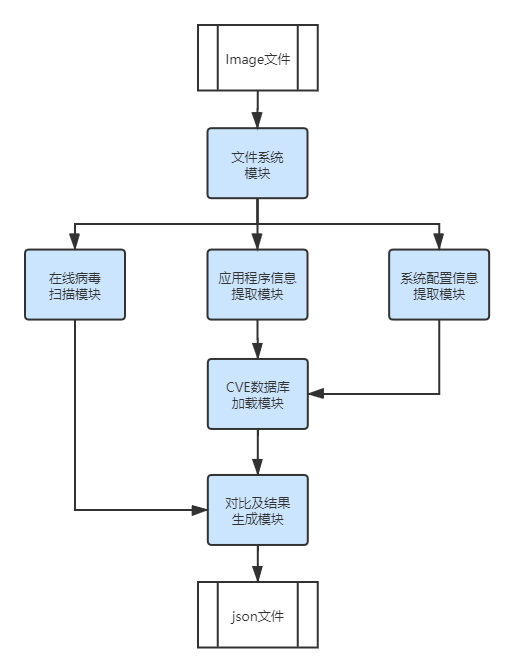


根据系统功能图，可以将本项目划分为六个模块，分别为：

1. 文件系统模块
2. 在线病毒扫描模块
3. 应用程序信息提取模块
4. 系统配置信息提取模块
5. CVE数据库加载模块
6. 对比及结果生成模块

### 2.3 子模块关系

以模块化的方式展示整个系统的流程图如下所示：



# 3 各模块详细设计

### 3.1 文件系统模块

#### 3.1.1 模块功能描述

【流程图】

本模块的功能

#### 3.1.2 镜像文件的解析

#### 3.1.3 文件系统的构建

###### 3.1.3.1 思路

###### 3.1.3.2 算法

* Xxx

#### 3.1.4 输入输出与性能分析

### 3.2 在线病毒扫描模块

#### 3.2.1 模块功能描述

【流程图】

本模块主要功能为

#### 3.2.2 在线病毒扫描

#### 3.2.3 算法

#### 3.2.4 输入输出与性能分析

（1）

### 3.3 应用程序信息提取模块

#### 3.3.1 模块功能描述

本模块的主要功能为

#### 3.3.2 应用程序信息介绍与提取

#### 3.3.3 算法

#### 3.3.4 输入输出与性能分析

### 3.4 系统配置信息提取模块

#### 3.4.1 模块功能描述

【流程图】

本模块的主要功能为

#### 3.4.2 系统配置信息的提取思路

#### 3.4.3 算法

#### 3.4.4 输入输出与性能分析

### 3.5 CVE数据库加载模块

#### 3.5.1 模块功能描述

【流程图】

本模块的主要功能为

#### 3.5.2 CVE数据库文件分析

#### 3.5.3 算法

（1）数据库的搭建

（2）CVE数据库的请求与更新

（3）

。。。

#### 3.5.4

### 3.6 对比及结果生成模块

#### 3.6.1 模块功能描述

【流程图】

本模块主要是辅助模块，是与用户交互的界面，询问/帮助用户如何使用程序、处理检测的结果。

#### 3.6.2 算法

（1）程

#### 3.6.3 输出结果

# 4 工具对比及分析

### 4.1 现有工具的缺陷

### 4.2 本工具的优势