哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY

2021 年秋季学期

计算学部《软件安全》

Lab5 实验报告

姓名	余涛，崔同发
学号	1180300829，1180300801
专业	信息安全
班号	1803202
手机号码	15586430583

1、实验项目描述

1、 理解基于异常检测的恶意攻击行为检测方法

（1） 掌握异常检测的流程

（2） 学习相关的异常检测算法

2、 基于距离的异常检测方法

（1） 掌握欧氏距离的概念

（2） 如何选取数据集的属性集

（3） 选取合适的检测模型

3、 基于 KD 树的网络流量异常检测模型

（1） 利用 KD 树构建一个用于多维空间最邻近搜索的数据结构

（2） 建立历史数据集合

（3） 规格化数据

（4） 采用标准分割策略，进行基于维度分割的 KD 树构建

（5） 基于待检测数据 X，利用构建好的 KD 树搜索，找出历史数据集合中与待检测数据 X 最近的数据；计算二者之间的欧氏距离，与阈值比对，确定是否是异常数据点。

2、实验要求

1、 实验数据准备。利用 KDD1999 数据集(KD 树异常检测.pdf p31 页-34 页)提供的数据进行实验。选取部分正常数据做为训练集，选择部分攻击数据和剩下的正常数据做为测试集。

2、 可以只选择流量属性集(KD 树异常检测.pdf p34 页)。可只针对 DOS(smurf 攻击即可)攻击进行异常检测，其他攻击不考虑。

3、2 人一组完成实验。

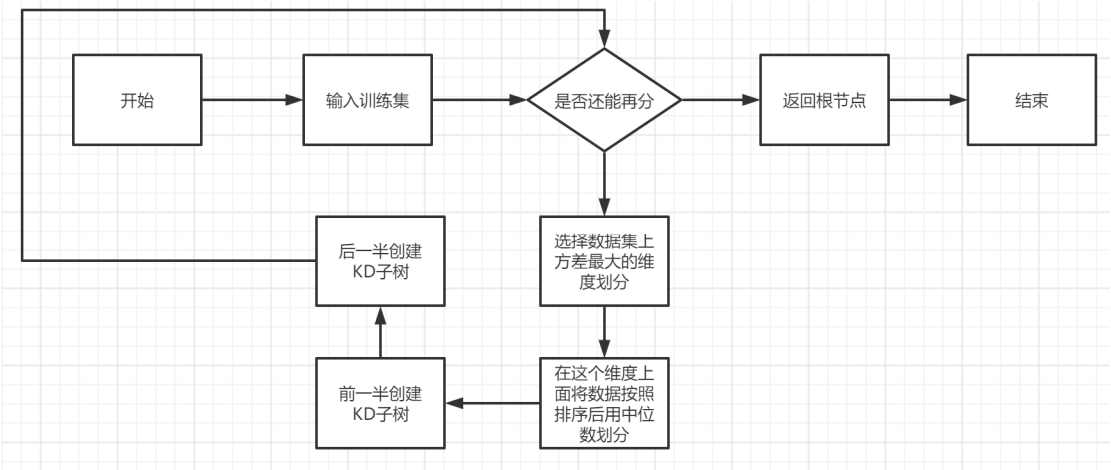
4、下载阅读 “实验 5 相关资料”中的 KD 树异常检测.pdf 文件

5、利用 “实验 5 相关资料”中的 kdd 原始数据中的 kddcup.data_10_percent 作为数据（看数据说明）。数据中每条都有标记为：NORMAL 或 ATTACK 类型。利用标记为 NORMAL 的数据建模（构建 KD 树）。利用一部分标记为 NORMAL 的

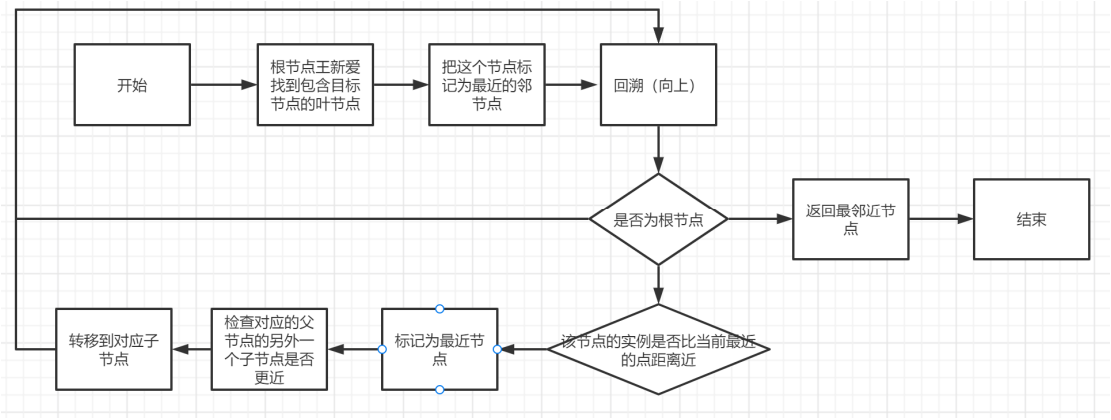
数据 和 ATTACK 数据作为测试数据。（数据集中可能存在少量有错误的数据，注意）

3、实验结果

1、创建 KD 树：



2、在 KD 树上面搜索最近邻



3、数据结构：

```
class KD_node:
    def __init__(self, point = None, split = None, left = None, right = None):
        self.point = point # 数据点的特征向量
        self.split = split # 切分的维度
        self.left = left # 左儿子
        self.right = right # 右儿子
```

4.实验结果：

```
C:\Users\10636\PycharmProjects\softsecutriy_lab5\venv\Scripts\python.exe C:/Users/10636/PycharmProjects/softsecutriy_lab5/lab5.py
299
500
现在开始运行, 训练集大小为: 600
检测的Normal数量是: 500
测试的Smurf数量是: 299
测试的Normal数量是: 500
总训练时间:0.5399518013000488s
总测试时间:4.395253419876099s
```

测试结果准确