|  |  |  |  |
| --- | --- | --- | --- |
| **论文开题报告进展表** | | | |
| **姓 名** | 胡希文 | 填表时间 | 2109年09月04日 |
| **论文题目** | 基于机器学习的入侵检测研究 | | |
| 已读文献1：DARPA Intrusion Detection Evaluation  文献来源：<https://archive.ll.mit.edu/ideval/docs/attackDB.html>  文献主要研究内容：介绍DARPA 1999数据集的攻击方式，一共四大类，53种攻击方式。四大类分别是拒绝服务攻击(Denial of Service Attacks, DoS)、未授权的本地root访问(User to Root Attacks, U2R)、未授权的远程访问(Remote to Local Attacks, R2L)、端口监视或扫描(Probes)，还有一类较少的攻击方式Data。  已读文献2：A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems  文献来源：Kris Kendall's MIT Master's thesis  文献主要研究内容：文献先是介绍了网络和入侵的相关知识，然后探讨了1998年DARPA数据集中出现的攻击方式，然后整理出了四大类入侵方式，共计32种，后世的入侵检测攻击方式很大程度上参考了该论文，然后该论文又介绍了7种在当时流行的攻击方式，分别是Cracker, Spy, Rootkit, Http Tunnel, SNMP Monitoring, Multihop, Disgruntled/Malicious User。  已读文献3：入侵检测系统的决策树生成方法  文献来源：河南工程学院学报 Vol. 31, No 2, Jun. 2019  文献主要研究内容：网络入侵检测的主要分为误用检测和异常检测，然后本文选取的检测方式为异常检测，选用的数据结构为决策树。该论文选取的是DARPA98数据集来学习和评估决策树，首先是走DARPA数据集中抽取数据，然后进行数据预处理，最后再用ID3算法生成决策树。  由于DARPA数据集中，将攻击划分为四类，故，本文中生成四个决策树。但是由于篇幅限制，该论文生成了U2R和R2L攻击的决策树。最后该论文对决策树方法在入侵检测中的准确率进行了实验评估。 | | | |
| **开题报告进展情况**： | | | |
| （本周关于开题报告所进行的相关工作）  1、了解了入侵检测系统的工作流程  2、了解了入侵检测系统的研究方法  3、了解了入侵检测系统常用的数据集，DARPA数据集，并对该数据集的四大类53种攻击方式做了详细的了解  （下周计划）  1、目前对机器/深度学习还不是很了解，下周的重点将会放在这一块  2、学习python，平时用java多点，几乎没用过python，尽快熟悉起来  3、如果有余力，搭建一个小的测试环境尝试一下 | | | |