丁雄

dingxiong@iie.ac.cn · (+86) 182-7193-1699 · 88daxiong

教育背景

中国科学院大学 (保送)，网络空间安全，工学硕士 2017 – 至今

获得荣誉：校三好学生，优秀学生干部，校优秀共产党员

华中科技大学，计算机科学与技术，工学学士 2013 – 2017

获得荣誉：国家励志奖学金 *(2 次)*，校级奖学金 *(3 次)*，校三好学生，优秀学生干部，优秀学生党员，校优秀毕业生

IT 技能

- 熟悉计算机理论，数据结构，操作系统，计算机网络以及基础算法。
- 熟练使用Python 完成开发任务，了解C、C++ 基本语法，可在Linux 环境下编程。
- 熟悉机器学习经典算法，例如支持向量机、决策树，最近邻等；熟悉深度学习中常用的神经网络类型，例如RNN、CNN 等；
- 熟悉数据库理论和SQL 语言，熟悉常用数据库，例如MySQL、Elasticsearch 等。
- 了解Map-Reduce 模型，以及Hadoop、Spark 等平台相关算法实现和架构设计。
- 熟练使用Git 进行版本控制和代码托管、Markdown 进行文档编写。

项目经历

基于机器学习的鱼叉式钓鱼邮件检测发现 2018 年 11 月 – 至今

技术栈：*Python* + 数据增强 + 随机森林/决策树/朴素贝叶斯/支持向量机

- 毕业设计项目，已投稿 CoopIS，正在申请专利。
- 从原始Pcap 包和邮件信息提取出相关特征，包括邮件转发关系特征、邮件信誉特征、邮件附件特征以及常用的邮件特征。使用SMOTE 算法进行数据增强。最后使用机器学习算法分类。
- 10000+ 样本量，识别准确率超过 99%。

基于深度学习的恶意代码家族识别系统 2018 年 10 月 – 2019 年 4 月

技术栈：*Python* + *MySQL* + *Keras* + 卷积神经网络

- 为某单位开发的恶意代码家族识别系统，已正式部署上线应用，正在申请软件著作权。
- 5 种预处理方式。分别是使用 IDA 提取出的控制流图 (CFG 图片)、切片矩阵 (TXT 文件)，直接提取二进制的灰度图 (灰度图片)，提取 PE 结构的直方图 (直方图 TXT 文件) 和导入链接库 (DLL 的 TXT 文件) 等。使用 CNN 进行训练和分类。
- 800 万样本量，模型的准确率超过 98%，能实现超过 50 种家族识别。

基于云平台的大规模恶意代码自动化分析系统 2018 年 9 月 – 2019 年 6 月

技术栈：*Python* + *MySQL* + *Elasticsearch* + 同源分析

- 为某单位开发大规模恶意代码自动化分析系统二期。
- 使用自研沙箱和市面上 Cuckoo 沙箱对恶意样本进行动态分析。提取出行为、网络、注册表和文件等行为，得到详细的分析报告。将报告解析后放入 Elasticsearch 中，利用其关联检索实现同源分析。
- 当前每天分析 2 万个样本，可随着服务器数量而扩展。

其他

- 技术博客: https://88daxiong.github.io/
- 语言: 英语 - 熟练 (CET-6)
- 性格: 积极向上，做事认真负责，团队协作意识强。