项目编号：

**大学生创新创业孵化项目**

**申 报 书**

**申请者姓名： 朱佳超**

**所 学 专 业： 计算机科学与技术**

**孵化项目名称： 软件代码漏洞智能检测系统**

**项目所属行业： 软件安全性测试**

**项目承担单位（公章）：杭州电子科技大学**

**填报日期： 2020 年 12 月 1 日**

**浙江省大学生科技创新活动计划暨新苗人才计划实施办公室 制**

填写说明

一、凡申报本孵化项目的申请人，须认真阅读本项目有关通知，封面申请者姓名一栏填写项目负责人。

二、格式要求：申报书中各项内容以Word文档格式填写，表格中的字体为小四号仿宋体，1.5倍行距；表格空间不足的，可以扩展或另附纸张；均用A4纸双面打印，于左侧装订成册。

三、申报书由所在学校领导审查、签署意见并加盖公章后，一式1份（原件），报送浙江省大学生科技创新活动计划暨新苗人才计划实施办公室。

1. **项目简介**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 项  目  概  况 | 项目名称 | | | 软件代码漏洞智能检测系统 | | | | | | | |
| 项目性质 | | | （）基础研究 （√）应用研究 | | | | | | | |
| 项目来源 | | | （）自主立题 （√）教师指导选题 | | | | | | | |
| 起止时间 | | | 自2020年 12月 至 2022年 12月 | | | | | | | |
| 项目状况 | | 1、研发阶段√ 2、中试阶段 3、批量（规模）生产（选项打√） | | | | | | | | | |
| 申请人 | 姓名 | 朱佳超 | 性别 | | 男 | 出生年月 | 1996.04 | | 入学年份 | 2019.09 | |
| 所在院系 | 计算机学院 | 学号 | | 192050147 | 联系电话 | 15869114920 | | 电子信箱 | 3232283602@qq.com | |
| 项 目 组  主要成员 | | 姓名 | 性别 | | 年龄 | 学院 | 专业 | 学号 | 具体分工 | | |
| 陶泓宇 | 男 | | 22 | 计算机 | 电子信息 | 202050226 | 前端界面、前后交互 | | |
| 武佳杰 | 男 | | 30 | 计算机 | 电子信息 | 204050014 | 漏洞检测、整体逻辑 | | |
|  |  | |  |  |  |  |  | | |
|  |  | |  |  |  |  |  | | |
| 项  目  指  导  教  师 | | 姓名 | | | 徐向华 | 性 别 | 男 | | 出生年月 | | 1965.10 |
| 主要研究方向 | | | 数据挖掘、工业互联网安全、软件代码测试、物联网协议优化设计 | | | | | | |
| 近三年获奖成果：国家级\_\_\_\_项，省部级\_\_2\_项 | | | | | | | | | |
| 近三年科研经费\_\_300\_\_\_\_万元，年均\_\_100\_\_\_\_万元 | | | | | | | | | |
| 项  目  主  要  内  容  简  介 | | 随着信息技术和软件技术地不断发展，应用于各种场景的软件也具备了更加全面丰富的功能。随之而来的是更多的代码量以及更多安全性和稳定性方面的问题。这类问题如果没有得到好的解决，极易给公司和用户带来潜在的安全风险，导致不必要的经济损失。  软件代码漏洞智能检测系统主要可以分为两部分，第一部分是基于深度学习的静态代码漏洞分析，第二部分是基于模糊测试的动态漏洞挖掘。  基于深度学习的静态代码漏洞分析。在之前的漏洞检测中，大部分的工作是由人工专家完成的，比如漏洞特征的定义，漏洞规则的改写等，这些检测工作非常依赖于专家的参与，如果通过专家定义的一些规则或者特征与当前的漏洞相关性比较小，那么一般很难检测出当前软件或者系统中的漏洞。而使用深度学习的方法，可以通过现有的一些漏洞数据，去训练某个深度学习模型（eg. LSTM、BLSTM、CNN等），之后通过模型自动学习出漏洞的特征。最后，使用已经训练好的模型对待检测的软件或者系统进行测试，以期找出其中的漏洞。  基于模糊测试的动态漏洞挖掘，主要是调用成熟的模糊测试框架如libfuzzer、AFL等对测试对象进行分布式的测试，但在功能上比现有框架更加全面，在测试效率上比现有框架更快，在操作上比现有框架更方便。软件代码漏洞智能检测系统能够检测出可能造成软件崩溃的一些崩坏点或者软件自身的BUG。在软件被改进之后能够即时显示BUG点的修复情况，提高软件系统的健壮性和安全防护能力。  我们可以利用深度学习技术去预测漏洞的发生以及定义漏洞的类型，之后与fuzzing技术相结合，让软件漏洞挖掘与测试更加可靠与高效。 | | | | | | | | | |

**二、项目的研究目的及意义**

|  |
| --- |
| 1、申请项目的必要性、目的及意义  **必要性**  随着技术的发展，各种应用程序、App应运而生，软件的各种漏洞随着软件的上线也会逐渐暴露出来。但只凭软件测试人员的人工测试和用户的反馈来修补漏洞是非常有限的。一个未发现的系统漏洞可能会带来无法估量的经济财产损失。  在当前经济全球化的大趋势下，软件作为商品流通在这个大环境之中。软件信息安全一直是软件领域的一个重点。如果软件的一些致命漏洞没有被检测出来，在全球化过程中被利用，造成得不仅仅是信息泄露、系统崩溃这种小问题，甚至会损害到国家利益，给国家造成严重损失，例如2020年4月份，由于软件系统中的某个漏洞被黑客利用，给以色列水利基础设施造成了严重的破坏，也严重影响了人们的日常生活。  **目的及意义**  本系统对软件进行全面的分析测试，检测软件中可能存在的漏洞以及漏洞的种类，并且精确定位漏洞的位置以及漏洞出现的深层次的原因。自动化的执行减少了多余的人力开销，降低了软件安全测试开发的周期，提高了软件的安全性。对于找到的漏洞，系统会生成漏洞条目报告进行展示，以便进行漏洞的定位和修复。本项目可以满足企业在开发中对于软件安全测试的需求，大大减少了软件的安全测试开发成本。 |
| 2、项目的背景、主要内容、技术水平及应用范围  **项目背景：**  近年来软件安全和应用安全的重要程度在与日俱增，尤其是工业行业，作为近年来地缘政治争夺和网络空间对抗的主战场，攻击事件频发、多发。仅今年上半年，就有包括以色列水利基础设施、伊朗重要港口朗沙希德·拉贾伊、瑞士铁路机车制造商、葡萄牙跨国能源公司在内的多个大型工业企业、政府机构遭到攻击。  因此，2020年8月，国务院发布了《新时期促进集成电路产业和软件产业高质量发展若干政策的通知》，鼓励软件企业执行软件质量、信息安全、开发管理等国家标准，提高软件质量，增强行业竞争力。而对于那些没有经过充分安全测试、有效消除缺陷的“自带漏洞”的软件产品，如果任其大量涌入市场、投入实际应用，必将对已初步成形的国家网络安全保障态势带来新的隐患。  目前国内大多数企业的软件测试主要注重功能测试、性能测试等实用性方面的测试，而对于安全方面的测试关注过少，一方面是安全测试成本过高，测试人员需要了解大量安全方面的知识，另一方面是漏洞挖掘的不确定性，可能投入很多人力资源也无法将所有漏洞找全，并且该过程非常耗时。  因此，开发一个软件代码漏洞智能检测系统是极其必要的。  **主要内容：**  本项目搭建软件代码漏洞智能检测系统，主要分为两部分，基于深度学习的静态代码漏洞分析与基于模糊测试的动态漏洞挖掘，可以通过静态代码分析与动态漏洞挖掘两种方式来检测软件中存在的漏洞。主要的功能模块一共有6个，包括DL(deep learning)漏洞检测模块、模糊测试模块、漏洞报告生成模块、漏洞修复验证模块、集群资源管理模块、用户模块。  DL(deep learning)漏洞检测模块：包括使用系统提供的数据或者用户自定义的数据生成深度学习预训练模型，对用户提交的目标代码进行快速漏洞检测，生成相关漏洞特征，对发现的漏洞生成解释文件。  模糊测试模块：集成了多种模糊测试框架，可以满足绝大部分的测试要求。对待测试的软件进行插装编译后，将其分配给空闲的模糊测试节点，进行动态漏洞挖掘。本系统采用分布式架构，进行并行化的漏洞挖掘，且各个测试子节点会同步测试信息，用于提升漏洞挖掘的效率。  漏洞报告生成模块：对已经发现的漏洞，生成相关的代码漏洞检测与分析报告，以可视化的方式展示在系统中，让使用者对漏洞有一个整体的观测。  漏洞修复验证模块：当软件工程师修复了软件中已发现的漏洞后，可以通过此模块验证是否漏洞已经完全修复。如果是，就将漏洞标记为已修复。  集群资源管理模块：本系统采用分布式架构，此功能模块用于管理各个资源节点，可以显示有多少可用节点资源，并且给任务分配资源用于漏洞分析与挖掘。  用户模块：包含用户登录、登出以及用户权限控制相关内容。    图1 系统核心模块流程  **技术水平：**   1. 使用BLSTM、GRU、Bert等深度学习中的前沿研究成果，自动生成代码漏洞特征，对代码漏洞进行分类处理。 2. 结合深度学习技术，使用Fuzzing对漏洞进行二次挖掘，将相关技术整合到一起。   （3） 基于分布式架构，改进并行化模糊测试算法，用于提升漏洞挖掘的效率。  **应用范围：**  本系统支持所有泛c类语言的开源应用软件。 |
| 1. 实施该项目所具备的基础、优势和风险   **基础：**  技术研发方面，项目团队成员有两年以上的开发编码经验，掌握网络威胁检测、数据收集、数据预处理、数据挖掘、漏洞分析、网络爬虫、前端平台搭建等相关技术，能够应用这些技术到本项目中。工程能力方面，本团队对软件漏洞测试项目已有2年多的研究基础，针对目标软件的源代码采用已有的模糊测试工具进行漏洞检测并修改已发现的漏洞。  指导老师方面，本项目指导老师现任浙江省数据存储传输技术重点实验室副主任，杭州电子科技大学智能与软件技术研究所副所长，工业互联网研究中心副主任，他主持国家973项目课题1项，国家科学基金面上项目2项，浙江省重大科技专项和杭州市科技计划项目6项以及横向课题多项，作为主要成员参与国家科技支撑计划项目、国家自然基金、浙江省重大科技专项等项目10多项，涉及研究方向包括工控安全、网络安全、物联网安全、数据挖掘等多个方面，具有丰富的项目经验。  实验室和杭州安恒信息、浙江乾冠信息安全研究院等国内知名网络安全公司开展了产学研合作工作。  承担相关项目：浙江省重点研发计划项目：基于数据挖掘的工业控制系统安全漏洞分析与态势感知技术研发（2017C01065），2017-2020，850万  已经申请网络安全相关发明专利6项：   1. CN 202010512591.6一种工业控制设备固件的模糊测试方法，徐向华，邵帅，朱佳超；杭州电子科技大学 2. CN 202010467332.6一种覆盖率引导的VxWorks内核模糊测试方法，徐向华，邵帅；杭州电子科技大学 3. CN201910660498.7基于流量追溯的工控协议模糊测试用例生成方法;申请日 : 2019.07.22公开（公告）号 : CN110401581A公开（公告）日 : 2019.11.01申请（专利权）人 : 杭州电子科技大学;发明人 : 徐向华; 邵帅; 王然; 4. CN201910615367.7监控基于流量重放的工控协议模糊测试方法;申请日 : 2019.07.09公开（公告）号 : CN110505111A公开（公告）日 : 2019.11.26申请（专利权）人 : 杭州电子科技大学;发明人 : 徐向华; 邵帅; 王然; 5. CN 201910302737.1一种基于稀疏自编码器的刷榜者检测方法，申请人：华东师范大学、杭州电子科技大学；发明人：何道敬、潘梦函、郑康锋、徐向华、唐宗力、潘志松、陈铭松 6. CN 201910392107.8、一种基于集成学习的群体口令强度评价方法，申请人：华东师范大学、杭州电子科技大学；发明人：何道敬、周贝贝、杨肖、徐向华   开发了工业互联网安全相关软著三项：   1. 2019SR0245113 工业控制系统漏洞挖掘检测软件V1.0，杭州电子科技大学 2. 2019SR0245119工业控制系统安全态势感知平台V1.0，杭州电子科技大学 3. 2019SR0781284 工业互联网安全态势感知平台V1.0，杭州电子科技大学   2020-06-01 23-00-47屏幕截图  图2 工业控制系统漏洞挖掘检测  2020-06-01 23-12-56屏幕截图  图3 工业控制系统漏洞挖掘检测结果  Selection_003  图4 工业互联网安全态势感知界面图  **优势：**  项目团队成员均有过项目开发经历，编程基础扎实，具有较强的学习能力。项目团队成员中有博士生，其主要的研究方向就是基于深度学习的软件漏洞分析，擅长多种深度学习框架，并已阅读了大量的相关论文。项目团队成员中还有参与了“基于数据挖掘的工业控制系统安全漏洞分析与态势感知技术研发”省重点研发项目的同学，了解项目的开发流程，对于漏洞安全挖掘有过基本的了解和学习，能够使用主流的模糊测试框架，并且跟踪阅读了最近几年模糊测试方向的文献资料。  本项目指导老师具有丰富的项目指导经验，能够帮助解决项目研发过程中遇到的技术和实施难题，保障项目顺利完成。  **风险：**  本项目实施过程中存在一些技术风险，具体如下：  （1）将深度学习用于软件漏洞分析是一个比较新的方向，所以该领域的相关文献资料还比较少。  （2）如何利用深度学习的漏洞分析结果来指导模糊测试过程还有待学习和研究。  （3）并行化的模糊测试会存在任务冲突和工作量不平衡的情况，需要阅读更多的文献来解决这些问题，以提升漏洞挖掘的效率。 |
| 4、项目计划目标  项目计划实现以下目标：   1. 集成多种深度学习模型，实现静态代码的漏洞分析。 2. 集成多种模糊测试框架，依据任务实现对指定的模糊器的调度，完成模糊测试。 3. 实现分布式架构，利用更多的计算节点来提升漏洞挖掘的效率。 4. 实现多节点之间的测试信息同步，提升模糊测试的效率。 5. 实现漏洞信息的显示与管理，工程人员可以依据漏洞信息来修改漏洞。 6. 实现对漏洞修复的验证，检测软件中漏洞是否已经修复。 7. 对目标代码进行测试，并修改bug，最后通过验收测试。 |

**三、预期成果、知识产权形成及经济、社会效益分析**

|  |
| --- |
| 5、 项目的预期成果及知识产权归属情况  软件代码漏洞智能检测系统：  申请发明专利2项；  软件著作权1项；  论文1篇；  本项目产生的知识产权归项目团队成员和指导老师共同拥有。 |
| 6、项目的市场前景分析  现在市场上存在很多的应用软件，而且这些软件都是经过用户使用后才显现出安全漏洞，此时程序员才会进行漏洞检测和修复。并且大多数依靠人工进行漏洞检测，而且人工检测到的漏洞基本很少，且效率低。本系统不仅通过自动化检测漏洞，而且采用多个模糊器同时对目标进行检测，发现漏洞并修复，使其正常运行。友好的web交互界面实现更高的可用性。本系统具有高效、低成本、实用性强、操作简单等特点，可以融入整个软件测试流程之中，所以市场前景乐观。 |
| 7、项目的盈利能力分析及财务预算  系统研发成功后，因本系统具有高效、低成本、实用性强、操作简单等特点，可以向监管机构及软件开发公司进行推广使用，进而产生经济效益。  本项目财务预算一共1万元，支出详细如下表：   |  |  |  | | --- | --- | --- | | 支出项 | 依据和理由 | 金额（万元） | | 设备材料费 | 购置处理器、内存条等配件 | 0.3 | | 差旅会议费 | 去本地或外地企业、高校进行学术调研和会议交流 | 0.4 | | 知识产权费 | 用于发表论文、申请专利和软件著作权 | 0.3 | | 总计 |  | 1 | |
| 8、项目的社会效益分析  本项目的社会效益主要表现在以下两个方面：   1. 对于国家，通过软件代码漏洞智能检测系统对国家级的应用软件进行安全漏洞检测与修复，保证国家的战略安全，减少国家安全信息的泄露，维护国家网络安全。 2. 对于社会，系统提高对各大应用软件的安全漏洞检测和修复，不仅降低信息泄露的风险，促进了社会稳定，而且也避免不法分子利用漏洞进行非法操作，提高社会生产效益。 |

**四、项目实施进度方案**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 本项目进度安排如下表所示：   |  |  | | --- | --- | | 起止年月 | 进度目标 | | 2020.12~2021.01 | 项目需求调研，关键技术研究 | | 2021.02~2021.04 | 总体方案设计 | | 2021.05~2021.11 | 系统模块开发、集成、测试 | | 2022.02 | 项目中期检查 | | 2022.02~2022.08 | 系统改进 | | 2022.09~2022.11 | 项目验收 | |

**五、项目组承诺**

|  |
| --- |
| **承 诺 书**  以上所填内容真实可靠，本项目组承诺：该项目立项后，将严格遵守有关规定、遵守本申报书和预算表中规定的条款和内容，保证按计划进度完成项目任务。  项目组全体成员（签章）：    年　　月　　日 |

**六、学校审核意见：**

|  |
| --- |
| 负责单位（公章）  年 月 日 |

**七、专家组审核意见**

|  |
| --- |
| 专家组组长签章：  年 月 日 |

**八、实施办公室审核意见**

|  |
| --- |
| 公章：  年 月 日 |