基于静态代码度量的安卓应用安全漏洞预测方法研究

一. 成果完成人及所在单位：王利欣，王素芳，计算机与信息工程学院　指导教师：崔建峰

二. 成果简介：

本研究分析了现有的风险脆弱性预测模型机器学习算法和不同的风险水平。可以得出结论从静态代码度量的实验结果很有可能用于预测安卓的风险漏洞各种物联网应用。这项研究的发现对基于安卓构建了一个高效的风险脆弱性预测模型实时场景中的物联网系统。本研究的主要局限性区域）数据集样本较小，以及（b）不平衡数据集。在未来，我们打算在大型数据集上进行实验，并对基于 ML 的安全措施的性能进行了度量对于安卓应用程序。优化射频参数减少外部因素对预测有效性的威胁模型。此外，还研究了其他方法在 ML 中的结合 Android 风险漏洞预测可以为物联网提供更好的结果应用。另外，在未来，我们期待着选择更多的静态代码度量来提高物联网应用预测模型。

三. 主要研究内容及创新点：

随着近年来通信技术的发展，利用物联网（IoT）实现节点间的互联互通正得到极大的发展。物联网系统一般包括传感器和执行器，它们积累详细的数据并完成相应的动作。多个节点的融合实现了许多领域的物联网，包括智能家居、智能交通、智能城市等。然而，物联网中多个节点的关联会影响各种安全问题的结果；因此，物联网系统需要大量的安全措施，仅允许授权实体访问数据。图 1 是多节点连接示意图。

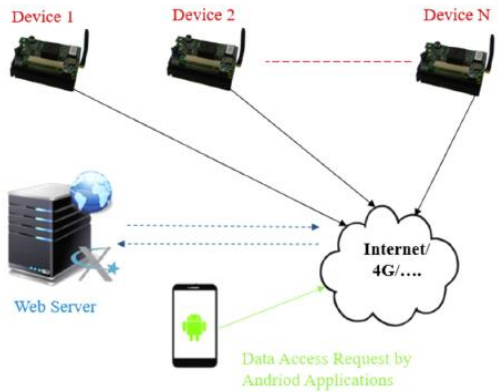


图 1 多节点互联

本研究使用 1406 个安卓 App 作为数据集，使用专业的分析工具 Androrisk 和 Sonarqube 对这些 App 进行了安全测评和特征提取。在本研究中共提取了 21 个静态特征，分别输入到 6 种不同的机器学习算法中，进行了研究。本研究的技术路线分别为数据预处理、特征提取与模型构建三个步骤。为了使这些分数可以用于分类问题中，我们使用 KNN 聚类算法，将这些风险分数划分为五个类型，随后，将这些 App 的源代码（Java 代码）使用 Sonarqube 工具进行了特征提取，分别提取了 21 个静态代码特征。使用这些特征，本研究构建了安卓漏洞预测模型。

四. 主要研究成果及应用情况

（1）论文一篇 Cui J, Wang L, Zhao X, et al. Towards predictive analysis of android vulnerability using statistical codes and machine learning for IoT applications[J]. Computer Communications, 2020. (SCI 三区已见刊)

联系人：王利欣，　　电　话：15602278321　　Email:1822021084@stu.xmut.edu.cn