目　录

摘　要 III
Abstract IV
1　绪论 1
1.1　研究背景 1
1.2　国内外研究现状 2
1.2.1　安卓恶意软件检测与家族分类研究现状 2
1.2.2　小样本学习研究现状 4
1.3　主要研究内容 6
1.4　章节安排 7
2　相关理论与技术 9
2.1　安卓系统介绍 9
2.1.1　安卓系统框架 9
2.1.2　安卓逆向工程 10
2.1.3　安卓安全机制 12
2.2　深度学习算法概述 13
2.2.1　卷积神经网络 13
2.2.2　长短期记忆网络 15
2.3　元学习 16
2.3.1　基于度量的元学习 16
2.3.2　基于优化的元学习 17
2.4　本章小结 18
3　基于动态多原型网络的小样本安卓恶意软件家族分类方法 19
3.1　总体流程 19
3.2　元学习策略 19
3.3　原型网络与多原型网络 20
3.4　方法的具体实现 22

3.4.1 API 调用的提取和特征向量化 23
3.4.2 将样本投影到特征空间的嵌入模块 24
3.4.3 动态多原型生成模块 25
3.4.4 Android 恶意软件家族分类模块 26
3.5 实验结果与分析 27
3.5.1 数据集介绍及划分 27
3.5.2 参数设置及评价指标 28
3.5.3 实验结果与分析 29
3.6 本章小结 36
4 基于增强原型网络的小样本安卓恶意软件家族分类方法 37
4.1 总体流程 37
4.2 方法的具体实现 38
4.2.1 加权原型生成模块 38
4.2.2 增强原型生成模块 39
4.2.3 家族分类模块 40
4.3 实验结果与分析 41
4.3.1 实验与参数设计 41
4.3.2 实验结果与分析 42
4.4 本章小结 48
5 总结与展望 51
5.1 总结 51
5.2 展望 51
参考文献 53
致 谢 59
攻读学位期间取得的学术成果和承担的科研项目清单 61

1 绪论

1.1 研究背景

随着移动互联网的快速发展，电子商务、医疗服务、游戏娱乐等应用市场逐渐由 PC 端转向移动端。由于智能设备的广泛适用性以及易于使用性，移动智能手机已经成为我们日常生活中的重要组成部分。近年来智能设备硬件和软件不断升级改进，推动着智能手机的使用和相关应用程序的快速增长，根据 GSMA 智库（GSMA Intelligence）的一份报告显示，到 2025 年，全球独立移动用户数量预计达到 59 亿，同年全球智能手机普及率将达到 80%[1]。在这些智能手机设备中，最受欢迎的操作系统是安卓。然而，由于安卓系统巨大的市场份额和其开源性，不仅吸引了合法开发应用程序的开发者，还吸引了传播恶意软件的攻击者。

在可观利润的推动下，安卓恶意软件出现了爆炸式增长，根据 AV-Test 对恶意软件的统计[2]，2023 年新的安卓恶意软件和潜在的有害恶意软件总数达到 309 万，主要类别有 PUA（Potentially Unwanted Applications）、Dialer、Trojan、Password 等。大量日益复杂的安卓恶意软件对移动设备用户构成了严重威胁，比如窃取用户隐私信息、推送垃圾广告、发送骚扰短信、恶意收费等。近年检测到的安卓恶意软件攻击事件中，一个名为柠檬集团（Lemon Group）的网络犯罪组织被发现利用名为“游击队”（Guerilla）的恶意软件感染了超过 50 个品牌的安卓设备，包括智能手机、手表和电视。Guerilla 恶意软件能够通过命令控制服务器，安装额外的恶意软件、从 SMS 拦截一次性密码、从感染的手机设置反向代理，允许攻击者利用受害者的网络资源、在受害者使用合法应用程序时向他们显示侵入性广告，这些功能使 Lemon Group 能够建立多样化的货币策略，从而实现巨额创收。此外，另一个著名的移动木马组织 Triada，也与这个 Lemon Group 网络犯罪组织有关联，Triada 组织持续对自己的恶意软件进行更新和优化，以满足各种不同的目标和环境需求。遥测数据显示，全球有数百万台受感染的设备正在运行，犯罪集团的恶意操作已遍布全球。因此，对不断发展的安卓恶意软件检测和分类的研究迫在眉睫。

研究者常用机器学习方法来对安卓恶意软件进行检测和分类，一些模型已经达到很高的检测效果。但是这些方法都是基于对大量样本数据进行分析，而新的恶意软件家族

种类在不断增加，且样本数量很小，要获取大量新的恶意样本十分困难。基于机器学习的模型很难对这些恶意样本进行有效的特征提取，导致检测和分类效率较低。因此，小样本下的安卓恶意软件家族分类成为了当前的重要研究方向。

1.2 国内外研究现状

1.2.1 安卓恶意软件检测与家族分类研究现状

目前常见的安卓恶意软件检测与分类方法主要有两大基本步骤，即提取数据特征和利用提取到的特征对分类器进行训练。对数据特征的提取主要采用静态分析方法和动态分析方法，提取完数据特征后，主要使用基于机器学习模型和深度学习模型两种方法进行检测和分类。

提取数据特征阶段根据是否需要在虚拟环境中实际运行安卓软件，分为静态分析方法和动态分析方法。静态分析不需要实际运行安卓软件，而是通过使用 Androwarn、TrIDNet 等静态分析工具或者 Androguard、Apktool 等反编译工具，从文件本身提供的内容中提取权限、API、意图、原始操作码等特征进行研究。该方法具有简单方便、检测速度快、节省时间和资源等优点；动态分析通过运行代码来分析安卓应用程序，通过跟踪执行阶段过程中的信息流，监测恶意软件的行为，从而收集有关应用程序的行为信息，如网络流量、API 调用序列、系统调用等。这种方法可以解决静态分析无法有效识别使用代码混淆、动态代码加载等技术的恶意软件的问题，但所需资源多、运行时检测速度慢、高计算开销等缺点也相对明显。

在提取完数据特征之后，进行一系列的数据处理与特征选择，最后通过训练分类器进行安卓恶意软件的检测与分类。传统的方法依赖于恶意代码特征库，并且特征库必须持续更新，以确保对新出现的恶意软件进行准确识别和分类。近年来，科研人员普遍利用机器学习（Machine Learning）与深度学习（Deep Learning）的方法，使用大量的数据特征训练模型，旨在有效识别恶意软件，并增强其对新型恶意软件的识别能力。

（1）基于机器学习的检测与分类

作为人工智能研究和应用的一个分支，机器学习运用各种算法对输入数据进行分析，并根据分析结果产生预测趋势或者分类等输出。近年来，许多研究都致力于利用机器学习技术来突破未知安卓恶意软件检测和家族分类的挑战。Bashir 等[3]提出了一种基于静态和动态分析的混合机器学习模型，采用支持向量机（Support Vector Machine，SVM）、

k-最近邻算法（k-Nearest Neighbor，KNN）、朴素贝叶斯（Naive Bayes，NB）等最先进的机器学习算法，对良性和恶性应用程序进行评估。Ünver 等[4]将安卓应用程序源文件转换为灰度图，从构建的灰度图中提取基于图像的局部特征和全局特征，用于训练多个机器学习分类器，包括随机森林（Random Forest，RF）、KNN、决策树、Bagging、AdaBoost 和 Gradient Boost。Liu 等[5]提出了一种基于本地函数调用图的安卓恶意软件多分类方法 NSDroid，该方法将安卓应用程序的子图结构抽象为许多节点的邻域签名，并将所有节点的邻域签名组合在一起，形成一个向量，用于表征整个应用程序的函数调用关系，最后将生成的签名向量输入到基于 SVM 的分类器中来进行家族分类。Onwuzurike 等[6]将 API 调用序列建模为马尔可夫链，利用状态转换的概率来构建应用程序的特征向量，利用机器学习算法 KNN、SVM 进行分类。Manzil 等[7]通过动态分析提取了安卓应用的系统调用，使用卡方技术进行特征选择，用于训练不同的机器学习分类模型，结果表明采用 SVM 和 AdaBoost 分类器准确率高达 99%。

传统的机器学习算法在处理简单和约束良好的分类和聚类问题时表现良好，但在面对复杂多变的恶意软件时显得力不从心。因此，一些研究者开始使用深度学习模型来识别和检测不断发展和增加的安卓恶意软件新变体。

（2）基于深度学习的检测与分类

深度学习[8]以其可以自动提取特征的能力而著称，这一特性消除了传统手工特征提取的繁琐过程，使其能够自动识别更为复杂且更具实用价值的高阶特征。He 等[9]提出了一种改进的 ResNeXt 模型，应用了基于卷积神经网络（Convolutional Neural Networks，CNN）的微调深度学习方法，在可视化的恶意软件样本上自动获取可区分正常数据和恶意数据的特征，通过嵌入新的正则化技术来提高分类任务的性能。Aldehim 等[10]提出了基于深度学习的 GBWODL-AMC 模型来进行 Android 恶意软件自动分类，模型使用基于 GBWO（Gauss-Mapping Black Widow Optimization）的特征选择方法来增强分类性能，GBWODL-AMC 采用深度极限学习机模型进行分类，其参数通过蚁狮优化算法进行优化选择，大量的实验结果表明，GBWODL-AMC 技术比其他恶意软件检测器具有更好的性能，最大准确率达到 98.95%。Aurangzeb 等[11]利用静态和动态混合分析提取特征，将特征作为输入提供给四个机器学习和一个深度学习分类器（Gradient Boosting，KNN，RF，XGBoost），然后应用集成学习的投票机制，根据多数投票系统将应用程序分类为恶意

软件或良性软件，解决了安卓恶意软件混淆变体的分类与检测问题。超凡等[12]使用静态分析提取特征，基于遗传算法进行特征选择，从而过滤掉无用特征，筛选出表达能力更强的特征，使用深度神经网络完成分类任务。

大量实验证明，深度学习模型正日益成为安卓恶意软件检测与分类的有效技术。然而，深度神经网络对样本数量要求比较高，样本量不足会影响检测和分类效果。

1.2.2 小样本学习研究现状

为了从有限数量的有监督信息的样本中学习，Bertinetto 等[13]提出了一种新的机器学习分支领域，称为小样本学习（Few-shot Learning），它主要解决了在机器学习场景中难以获取高质量数据集的问题。近十年来，对小样本学习的研究得到了广泛的开展，并取得了重大的研究进展，如阿里巴巴提出的 KGBert[14]在小样本学习领域的性能首次超过了人类。目前，针对小样本学习的有效方法主要包括数据增强、迁移学习、元学习和多模态学习：

（1）基于数据增强的方法

数据增强是指通过对现有数据进行颜色变换、几何变换、增加噪声、数据合成、混合样本和剪切等技术来生成更多的训练样本，以帮助模型更好地泛化和提高性能，可以有效减轻数据稀缺带来的问题。Gong 等[15]设计了一种名为 quad-patch 增强的新数据增强策略。利用样本图像的特征来分块和重新组合任何现有数据，从而生成伪新数据，丰富训练集。quad-patch 数据增强策略结合提出的双路径聚合注意力网络，从数据和架构两个方面解决小样本场景分类问题。Yao 等[16]基于几种常见的图像处理操作构建了数据增强操作池，以丰富数据增强空间，在此基础上，对数据叠加了各种操作，能够在保持原始输入图像核心特征的同时生成更多样的增广变量。实验结果表明，该方法在小样本目标检测方面优于其他方法。Zhang 等[17]利用预训练语言模型中的知识，构建 Cloze-style 式数据增强模型，采用无监督学习来强制增强数据在语义上与初始输入句子的相似度，并采用对比学习来增强每个类别的唯一性，在 CLINC-150 和 BANKING-77 数据集上验证了所提方法的有效性。

（2）基于迁移学习的方法

在迁移学习中，模型在源域的源任务上进行训练，源任务中有足够的训练数据，随后，经过预训练的模型被进一步引入目标域的目标任务中，进行再次训练和微调，以将

源任务中所学得的先验知识有效地迁移至目标任务领域。Yan 等[18]使用预训练的深度网络对输入和查询图像进行特征嵌入，对所得到的特征进行中心化和 L2 归一化，并使用欧氏距离作为最近邻分类的距离度量。Wu 等[19]提出了一种主动迁移学习算法 ACTrAda-TLSVM，该算法基于支持向量机，适用于跨领域数据分类场景。首先，利用最大均值差异计算源域样本权重的向量，并利用主动学习对源域中的样本进行标记，以获取高质量的标签训练样本；然后利用源域的知识和目标域的数据构建迁移学习分类器 TLSVM；最终将 TLSVM 和 TrAdaBoost 结合在一起构建 ACTrAda-TLSVM 分类器。Imtiaz 等[20]提出了一种适用于小样本任务的传导式拉普拉斯正则化推断方法，使用从基类中学习到的特征嵌入，最小化包含两个项的二次二进制分配函数：一个一元项将查询样本分配给最近的类原型，以及一个成对的拉普拉斯项支持附近的查询样本具有一致的标签分配。传导式拉普拉斯正则化推断方法不重新训练基础模型，也可以被视为对查询集进行图聚类，受支撑集的监督约束。但是当领域相关性相对罕见或大型辅助数据集不可用时，迁移学习有一定的局限性。

（3）基于元学习的方法

元学习即学会学习，是目前解决小样本学习的主流方法，其核心思想是让模型具备学会学习的能力。在这种范式下，模型被训练成拥有快速适应新任务或新环境的能力，而不是专门针对特定任务进行优化。MAML[21]是最早和最流行的基于优化的元学习方法之一，MAML（Model-Agnostic Meta-Learning）训练一组初始化参数，通过初始化参数 θ，该模型只需要较少的数据即可达到快速收敛。Chen 等[22]提出了 Evo-MAML，这是一种基于优化的元学习方法，在内循环中融入了进化梯度，避免了二阶信息，从而降低了计算复杂度。Pang 等[23]提出了 Adaptive-MAML，它由改进的 MAML 框架和神经网络模型组成，使用基于元增强的方法，可以在训练过程中自动生成虚拟样本，以缓解样本缺少并克服过度拟合问题。Zhang 等[24]将元训练中的梯度下降与元测试中基于度量的方法相结合，提出了一种新的元学习模型 GMT2（Gradient-based Meta-train with Metric-based Meta-test），建立了基于度量策略和梯度下降在元测试中的近似关系，通过直接计算相似度来对数据进行分类，训练好的元模型避免了收敛问题。

（4）基于多模态学习的方法

多模态学习更接近于人类智力的真实世界，它不再依赖于有限的样本，而是试图找

到其他模态的空间来帮助小样本学习。多模态学习整合了不同维度的信息，如语言、图像和音频。Pahde 等[25]利用生成对抗网络作为数据生成器对模型进行训练，基于语义信息生成相应的视觉特征，通过结合原始视觉特征获得增强的视觉特征，以此生成额外的训练图像，用于完成小样本视觉分类任务。Pan 等[26]提出一种多模态变分对比学习框架，该方法首先通过在视觉和语义空间中进行有监督的对比学习来训练一个变分自编码器，训练好的模型被用于从学习到的语义分布中采样特征，为支撑集增加样本，并为查询集生成伪语义，以实现支撑集和查询集中样本之间的信息平衡。并且该方法建立了一个多模态实例到分布的模型，通过变分推断将实例级别的多模态特征转化为分布级别的表示，从而促进了稳健的度量。多模态学习有望打破现实人类信息世界中小样本学习有用信息不足的困境。

1.3 主要研究内容

传统的安卓恶意软件分类依赖于大量已知的恶意软件，收集大量新检测到的恶意软件家族来训练分类器是极其困难的，而收集恶意软件样本又是一项费时费力的工作；其次，对于未知的安卓恶意软件类别或现有的新的恶意软件，常见的分类方法往往由于严重的过拟合而不能很好地工作。针对以上问题，本文的研究内容主要如下：

（1）针对稀缺样本问题和难以识别新家族的问题，本文采用了基于元学习的小样本学习技术，提出了基于动态多原型网络的小样本安卓恶意软件家族分类方法。首先，使用样本的特征嵌入向量，基于元学习中基于度量的原型网络方法，为各个家族生成初始原型；然后，针对同一类中样本分布成多模态分布问题，为每个家族生成多个原型；进一步，基于高斯混合模型的概率计算方法，对生成的多原型进行调整分配，将高斯分布的均值看作最终的原型，进而执行家族分类任务。在进行家族分类过程中，采用元学习方法不断的对分类器进行训练，并通过计算待分类样本与各个家族原型之间的距离来实现对恶意软件的家族分类。对于未能成功匹配的恶意软件，将其标记为一个未知的新家族，并从中提取家族特征生成原型加入到原型簇中，从而无需对分类器进行再次训练。

（2）针对来自同一家族的恶意软件样本对家族的代表性具有很大差异，而原型网络将每个样本的代表性看作相同的问题，为了进一步生成更具代表性的类原型，提出了一种类似于注意力机制的策略。该策略计算样本在家族中的重要性，为重要性大的样本赋予更大的权重，为重要性小的样本赋予更小的权重，利用加权和的方式得到家族的加

权原型。

（3）分类中的一个挑战是更紧密地映射同一类的样本，并更远地映射不同类别的样本。为了减少类内差异，同时扩大类间差异，提出了一种基于增强原型的安卓恶意软件家族分类方法。首先使用样本的特征嵌入向量，基于上述类似于注意力机制的方法生成家族的加权原型表示；其次，基于一个轻量级网络计算查询样本与家族原型之间的相关性，并学习一个适用于这两个嵌入之间的激活函数，利用激活函数对两个嵌入进行激活得到查询样本新的特征表示和家族的增强原型；最后，通过距离度量预测查询样本的标签来进行家族分类。在模型训练阶段采用负对数概率来定义分类损失函数，所提方法可以通过最小化查询样本上的分类损失来训练模型。

1.4 章节安排

基于以上研究内容，本论文共分为五个章节，各章节的内容安排如下：

第 1 章 绪论。主要介绍了本文的研究背景，说明了在样本数量较少的情况下进行安卓恶意软件家族分类的重要性，进而梳理了安卓恶意软件检测与家族分类和小样本学习的国内外研究现状，最后对本文的主要研究内容和章节安排进行了阐述。

第 2 章 相关理论与技术概述。主要介绍了安卓系统的体系结构、逆向工程和安全机制，同时介绍深度学习中的卷积神经网络 CNN 和长短期记忆网络 LSTM，并且对小样本学习中的元学习算法进行了阐述。

第 3 章 详细介绍了基于动态多原型网络的小样本安卓恶意软件家族分类方法。首先介绍了方法的总体流程与元学习策略，对单原型与多原型的计算进行了分析介绍；然后对所提出的方法进行具体介绍；最后在数据集上对所提方法进行了实验并与其他恶意软件家族分类方法进行比较，实验结果表明了所提方法的有效性。

第4章 详细介绍了基于增强原型网络的小样本安卓恶意软件家族分类方法。首先，介绍了方法的总体流程；其次，详细阐述了方法中两个重要的模块：加权原型和增强原型的生成；最后设计实验并与其他方法进行对比，验证方法的有效性。

第 5 章 总结与展望。对本论文所做的工作进行总结，并对未来进一步研究工作进行了展望。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/616045135011011005