Journal of Computer Applications 计算机应用 | ISSN 1001-9081 CODEN JYIIDU | 2019-08-21 http://www.joca.cn
doi:10.11772/j.issn.1001-9081.2019040730

# 基于互信息加权集成迁移学习的入侵检测方法

胡健[1*]，苏永东[1]，黄文载[1]，肖鹏[1]，刘玉婷[1]，杨本富[2]
(1. 云南电网有限责任公司 信息中心，云南省 昆明市 650217；
2. 云南云电同方科技有限公司，云南省 昆明市 650217)
(*通信作者电子邮箱 hujian92@126.com)

**关键词:** 入侵检测系统已成为网络安全体系结构中的必要组成部分。而现有的入侵检测方法的可行性和持续性，在面对现代网络安全需求时，仍然是一个巨大的挑战，主要体现在如何尽早地发现入侵威胁和提高入侵检测系统的检测精准度。本文提出一种基于互信息加权的集成迁移学习（Ensemble Transfer Learning，ETL）入侵检测模型（ETL based Intrusion Detection, ETLID），该模型首先通过迁移策略对多组特征集进行建模，然后使用互信息度量在迁移模型下特征集在不同域中的数据分布，根据度量值对多个迁移模型进行集成加权，得到集成迁移模型，该方法通过学习少量新环境下的有标记样本和以往环境下的大量的有标记样本的知识，可以建立效果优于传统非集成、非迁移的入侵检测模型，本文使用基准 NSL-KDD 数据集对该方法进行评估，实验表明，ETLID 具有良好的收敛性能、提高了入侵检测的精准率。

**关键词:** 入侵检测；迁移学习；互信息；集成学习；加权集成
**中图分类号:TP393.08** **文献标志码：A**

# Intrusion detection method based on mutual information ensemble transfer learning

HU Jian[1*], SU Yongdong[1], HUANG Wenzai[1], XIAO Peng[1], LIU Yuting[1], YANG Benfu[2]
(1.Yunnan Power Grid Co., Ltd. Information Center, Kunming 650217, China;
2.Yunnan Yundian Tongfang Technology Co.Ltd, Kunming 650217, China)

**Abstract:** Intrusion detection system has become an essential part of network security system, but there are still huge practicability and durability challenges for the existing intrusion detection methods, mainly reflecting on how to detect intrusion threats earlier and improve the detection accuracy of intrusion detection systems. This paper propose an ensemble transfer learning (ETL) intrusion detection method via weighted mutual information namely Ensemble Transfer Learning-based Intrusion Detection(ETLID). This model firstly separate the original source domain and the target domain data into several paired sub datasets and train several simple transfer learning models, then the model uses the mutual information to evaluates the relationships between the paired sub datasets from different domains, as well as calculates the weights of each simple model. Finally, the ensemble transfer learning-based intrusion detection system can be described with a linear weighting function, which weights the several models more efficiently. In the practical case that only few instances in the target domain are labeled and all instances from the source domain are labeled, an ensemble transfer learning model can be learned with only little information from the new environment and much knowledge from the prior environment. We use the benchmark NSL_KDD dataset to evaluate the ETLID and the results show that the ETLID method can reach higher precision and recall rate than the traditional intrusion detection system. Thus, the experiment results show that this ensemble method is superior to traditional non-ensemble methods and non-transfer methods

**Keywords:** intrusion detection**;** transfer learning**;** mutual information**;** ensemble learning; transfer ensemble

## 0 引言

新形的网络攻击呈现出了规模化、分布化、复杂化趋势，对入侵检测方法的有效性和及时性提出了更高的要求。目前普遍应用的以异常检测和误用检测（也叫基于签名的检测）

收稿日期: 2019-04-28; 修回日期:2019-07-22; 录用日期: 2018-08-05。
胡健(1992—)，男，云南文山人，工程师，硕士研究生，主要研究方向：信息安全、机器学习；苏永东(1967—)，女，北京市人，高级工程师，主要研究方向：信息安全；黄文载(1963—)，男，云南昆明人，高级工程师，主要研究方向：电力系统自动化；肖鹏(1988—)，男，云南昆明人，工程师，主要研究方向：网络空间安全；刘玉婷(1987—)，女，云南昭通人，工程师，硕士研究生，主要研究方向：信息安全；杨本富(1982—)，男，云南保山人，工程师，主要研究方向：软件工程、信息安全。

为代表的入侵检测技术普遍存在检测率低、误报过高以及过渡依赖知识库等不足。现有入侵检测方法发展已经遇到瓶颈，主要有三个限制性因素，第一是网络数据量的急剧增长，并将长期高速增长，需要快速在海量的网络流量中分析网络行为。第二网络应用的更高级更抽象，需要更加详细和丰富的上下文知识，提高入侵检测方法的监控深度和分析粒度。第三是网络协议的多样性和攻击行为的高级可持续性，增加了建立规范的难度。随着人工智能技术的快速发展，基于人工智能技术的入侵检测方法已成为入侵检测系统（Intrusion Detection System, IDS）[1]研究的热点之一。

## 1 相关研究

### 1.1 基于机器学习的入侵检测及其仍然存在的问题

将机器学习技术应用在入侵检测[2, 3]是入侵检测问题中的热点研究领域之一，其依赖于大量的有标注的网络访问数据，通过监督式学习方法（如决策树、支持向量机、神经网络等）对数据进行模式识别并建立分类模型，最后，借助该分类模型，对未来的网络访问实例进行判断，预测新的访问实例是否安全。图 1 是基于机器学习的入侵检测模型基本流程图，首先，研究者从数据仓库将历史访问数据取出，每一条历史访问数据包含了其在访问时候的一些附加信息，如访问时长、使用的是 TCP（Transmission Control Protocol）协议或是 UDP（User Datagram Protocol）协议等；然后，对这些数据进行标记，标记哪些是正常的访问，哪些是非正常的访问；最后，将这些数据作为训练数据置于机器学习算法中，通过监督式机器学习算法的训练，可以得到入侵检测的分类模型，并对未知的访问进行预测。

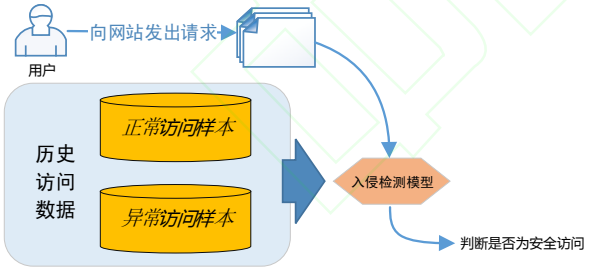


图 1 基于机器学习的入侵检测模型的基本流程

Figure 1 Basic flow of intrusion detection model based on machine learning

但是传统的机器学习算法在解决入侵检测时也有其弊端：

a. 传统的监督式机器学习算法基于两个基本假设：1. 训练数据的样本量足够多；2. 训练数据和真实环境的数据的分布相同。其中，前者保证了训练的模型足够可信，后者保证了训练的模型在新的环境下可用。但是，实际场景下，训练过程中使用的数据往往和真实环境下的数据存在数据分布差异；而给真实环境下的数据标记又是一件费时费力的工作，如何对缺乏足够训练样本的真实环境下的数据进行建模，是一项具有实际意义的工作。

b. 传统的机器学习算法应用在入侵检测系统上也有其局限性，研究者通常使用单一的机器学习算法对入侵检测数据进行建模，比如，仅仅使用神经网络[4, 5]或者支持向量机[6, 7]等来训练入侵检测的分类模型，虽然已经有研究者证明，对于经典的入侵检测数据集 KDD99 或者 NSL-KDD 而言，这些强分类器（如神经网络和支持向量机）已经取得了不错的效果[8]。但更新的研究也证明集成学习对于入侵检测是有利的[9]。

因此，如何在目标领域的数据量不足的情况下，使用集成策略（通过多分类器对各自适合的特征）进行模型训练并对各个模型进行有效集成，是解决实际的新环境下入侵检测问题的难点。

### 1.2 迁移学习和集成学习

为了弥补训练样本不足的问题，机器学习研究领域的很多研究者开始将目光投向迁移学习领域，迁移学习旨在能够学习相关领域 Ds（source domain，源域）的知识，并将之应用在另外一个数据分布不同但是却相关的领域 Dt（target domain，目标域），杨强博士早在 2009 年对迁移学习研究领域目前的研究进展进行了归纳，在其综述中，描述了基于样本的迁移、基于特征表达的迁移、基于关系的迁移、基于知识的迁移四种基本的迁移方式：其中，基于样本的迁移学习方法是在源域有标记样本充足、目标域有标记样本数据量很少的情况下，使用源域的有标记样本来辅助目标域构建模型的方法。由于基于样本的迁移学习算法易于实现，且与产业界的实际应用场景密切相关，目前已经在具体的领域，如银行的用户信用评估[10]、垃圾文本内容分类[11]、新闻文本分类[12]、推荐系统[13]、图片分类[14]等任务中被广泛应用。

另外，在具体的数据建模问题中，数据的来源具有多样化的特点，这也就导致了数据的各个特征可能分别属于不同的数据类型。如果使用 TrAdaboost[15]方法仅仅对某些特定类型的征建模，并不会有效利用好所有的特征信息。考虑到对不同类型的数据特征而言，有适合它的机器学习算法。因此，借鉴集成学习的策略，首先对不同的特征集（同数据集某几个特征组成的特征集合）独立进行迁移模型的训练，并在最后对这些模型进行有效组合，是可以提升迁移模型的效果的。

## 2 基于集成迁移学习技术的入侵检测

### 2.1 简单迁移模型策略

对于某特定的特征集进行迁移模型的训练，可以使用极简的迁移策略[16]，如图 2。

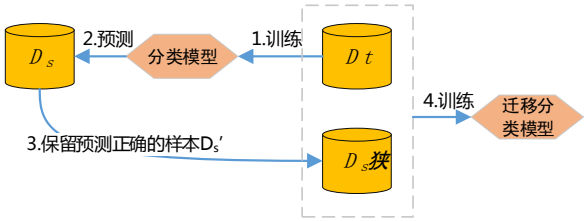


图 2 简单迁移模型的训练流程

Figure 2 Training process for simple transfer model

其中，$D_t$ 是少量有类别标签的目标域样本，$D_s$ 是大量的有类别标签的源域样本，通过极少的 $D_t$ 结合一种分类算法，可以训练得到一个简单分类模型，随后，将之应用在源域 $D_s$ 进行预测，保留预测正确的样本集 $D_s$'。最后通过混合 $D_s$'和 $D_t$ 的样本，使用相同的分类算法，训练得到迁移模型。这个模型有其优点：速度快，且适用于多个简单迁移模型集成。

## 2.2 集成迁移模型

在简单迁移模型的基础上，引入集成学习的概念，可得到集成迁移学习模型。首先，源域和目标域被成对地划分为不同的特征集，研究者可以在每对特征集上训练一个简单的迁移分类模型，与此同时，计算该特征集下的源域和目标域的数据分布的互信息值，并用这个互信息值来衡量不同的域在不同特征集上的差异情况；最后通过互信息值对多个简单迁移模型加权，得到最后的加权后的集成迁移模型。

实际上，对不同的特征集分别训练不同模型并将其组合的策略，在机器学习的相关研究中已经得到了充分肯定，比如在推荐系统中，宽深模型[17]就是最为经典的且具有高准确率的集成模型算法，宽深模型通过利用一个深度神经网络着重对连续型随机变量进行建模，然后使用逻辑回归模型对离散型随机变量进行建模，最后，通过再一层的逻辑回归模型学习到加权方案，就得到一个集成了“宽”模型和“深”模型的算法模型，原实验证明这种思路可以极大地利用好各类数据特征。

集成迁移学习模型的模型流程如图 2 所示：首先源域（$D_s$）和目标域（$D_t$）中的有标记样本按照不同的特征集被分为多组子源域（$D_{s1}$~$D_{sn}$）和子目标域（$D_{t1}$~$D_{tn}$），每组子源域 $D_{si}$ 和子目标域 $D_{ti}$ 可以训练得到一组简单的单模型迁移策略 $M_i$，同时，在迁移过程中，计算子源域和子目标域之间的互信息值并以之作为权重值来衡量模型的重要性，通过加权组合多个迁移策略（$M_1$~$M_n$），就得到了最终的集成迁移模型。

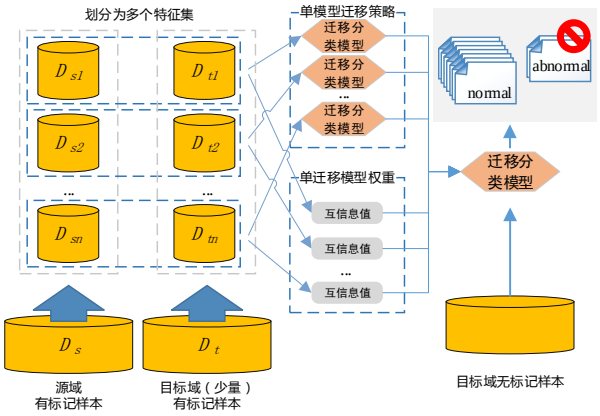


图 3 集成迁移学习模型

Figure 3 ensemble transfer learning model

## 2.3 加权集成方式

集成迁移模型的最终目的是将多个不同的迁移分类模型进行融合，而融合多模型最常用的方法是对多个模型的效果进行线性加权。加权的本质是对学习到的不同内容赋予不同的重要性，对于迁移学习而言，学习到的知识越有利于迁移，这个权值就应该越大。在集成迁移模型中，可以在不同特征集下使用源域和目标域的数据分布相似程度来评价简单迁移模型的迁移效果。

互信息（mutual-information）是用在信号学中的一个度量方法，用以衡量信号传输前后的损失或者差异，Ambusaidi 等人曾在 IDS 系统中使用了互信息[18]用来辅助特征选择，而在集成迁移模型中，通过互信息可以衡量由源域到目标域的分布差异。

当给定两个连续型随机变量 $U = \{u1, u2, \cdots, ud\}$和 $V = \{v1, v2, \cdots, vd\}$，其中 $d$ 表示样本个数，$U$ 和 $V$ 互信息之间的计算方式如下：

$$I(U;V)=H(U)+H(V)-H(U,V) \tag{1}$$

其中 $H(U)$和 $H(V)$分别表示随机变量 $U$ 和 $V$ 的信息熵，$H(U, V)$为 $U$ 和 $V$ 的交叉熵。

当 $U$ 和 $V$ 是连续(continues)型变量时，$U$ 和 $V$ 之间的交叉熵记为：

$$I(U;V)=\int_u\int_v p(u,v)\log\frac{p(u,v)}{p(u)p(v)}dudv \tag{2}$$

当 $U$ 和 $V$ 是离散(discrete)型变量时，$U$ 和 $V$ 之间的交叉熵记为：

$$I(U;V)=\sum_{u\in U}\sum_{v\in V} p(u,v)\log\frac{p(u,v)}{p(u)p(v)} \tag{3}$$

结合上面互信息的定义，定义使用互信息加权的广义集成方案如（4），其中 $M$ 是各个独立的迁移模型（$Mi$）的加权组合模型：

$$M=\frac{\sum_{i=1}^{n} I(U_i;V_i)\times M_i}{\sum_{i=1}^{n} I(U_i;V_i)} \tag{4}$$

## 3 实 验

### 3.1 入侵检测数据集

实验使用 NSL-KDD[19,20]基准数据集，该数据集是目前主要用于判断入侵检测系统性能的标准数据。NSL-KDD 数据集是 1999 年 KDD CUP 竞赛所使用的入侵检测数据集的改进版本，其包含了 41 个特征，训练数据集包含了 23 个类别标签，这 23 个类别标签隶属 5 个大类，而这 5 个大类中有 4 个大类（u2r、dos、r2l 和 probe）属于非正常的网络访问类别，正常的访问标记，只包含 normal 一种。训练数据集的类别标签的关系如表 1 所示。

表 1 NSL-KDD 数据中的入侵属性基本信息

Tab. 1 Basic information of intrusion attributes in nsl-kdd data sets

| 是否正常 | 访问属性 | 具体行为 |
|---|---|---|
| 正常 | normal | normal |
| 非正常 | u2r<br>dos<br>r2l<br>probe | normal,neptune,warezclient,ipsweep,portsweep,teardrop,nmap,satan,smurf,pod,back,guess_passwd,ftp_write,multihop,rootkit,buffer_overflow,imap,warezmaster,phf,land,loadmodule,spy |

通常情况下，研究者会将其中的 KDDTrain 或者 KDDTrain+_20Percent 作为模型训练的训练集，将 KDDTest 作为模型的测试集。而由于训练数据和测试数据的数据分布（均值和方差）存在差异，如表 2，且测试集中甚至出现了训练集中不存在的非正常网络链路的标记，如：saint, xsnoop, mailbomb, udpstorm, httptunnel, sendmail, sqlattack, worm, snmpguess, perl, mscan, apache2, xterm, named, snmpgetattack, processtable, ps, xlock，因此训练集和测试集的边缘概率分布和条件概率分布都不一致[21]，符合迁移学习的使用条件，因此该训练集和测试集也可以被认为是入侵检测领域的源域和目标域：

表 2 训练集和测试集数据分布差异

Tab. 2 The data distribution differences between the training dataset and the test dataset

| | 训练集 | | 测试集 | |
|---|---|---|---|---|
| 特征名称 | 均值 | 标准差 | 均值 | 标准差 |
| dst_host_rerror_rate | 0.12 | 0.31 | 0.23 | 0.39 |
| hot | 0.20 | 2.15 | 0.11 | 0.93 |
| rerror_rate | 0.12 | 0.32 | 0.24 | 0.42 |
| same_srv_rate | 0.66 | 0.44 | 0.74 | 0.41 |
| serror_rate | 0.29 | 0.45 | 0.10 | 0.30 |
| srv_rerror_rate | 0.12 | 0.32 | 0.24 | 0.42 |
| srv_serror_rate | 0.28 | 0.45 | 0.10 | 0.30 |
| wrong_fragment | 0.02 | 0.26 | 0.01 | 0.14 |

### 3.2 实验设置

首先，实验任务是一个迁移学习场景下的二分类任务，即在目标域的有标记样本量不足的情况下，结合迁移学习策略和集成学习方法，根据网络链路行为判断新的访问行为是否为入侵行为。

- **目标域数据划分**

由于在基于实例的迁移学习中，目标域难以获取足量的有标记样本，因此在实验中，只设置 0.1%~2.5%的目标域有标记样本来进行迁移学习模型的训练，而源域中的样本由于都是有标记的，因此可以完全被利用。

- **加权集成方案**

对于入侵检测的数据而言，数据往往包含了多种多样的数据类型：在 NSL-KDD 数据集中，其包含了分类特征和数值特征，而数值特征甚至也包含了已经被归一化的取值范围为[0, 1]的数值特征。实验拟对特征进行拆分，独立训练多个迁移学习模型并对模型的学习结果进行集成，使之更加有效利用原始数据的各类数据信息，使源域的知识能够最大化地被迁移到目标域中。

当使用数值型（记为 numr）特征、归一化数值（记为 norm）特征以及离散型（记为 cate）特征分别得到了多个模型 $M_{numr}$、$M_{norm}$、$M_{cate}$ 以及互信息值 $I_{numr}$、$I_{norm}$、$I_{cate}$ 时，使用互信息值对不同的预测模型结果进行加权集成，即得到了最终的学习器 $M$，记为（5）：

$$M=\frac{I_{numr}\times M_{numr}+I_{norm}\times M_{norm}+I_{cate}\times M_{cate}}{I_{numr}+I_{norm}+I_{cate}} \tag{5}$$

- **分类算法和评估方法**

在使用集成迁移模型来解决不同环境下的入侵检测问题中：实验使用多层感知机对取值区间为[0,1]的连续特征集建模，使用决策树对离散特征和取值区间为实数的连续特征建模。对于使用了迁移策略的实验而言，最终使用互信息加权策略对模型进行加权融合，得到最终的模型。

考虑到入侵检测的标准是既需要较高的精准率也需要较高的的的召回率，F1-score 作为一种常用的评估方法，兼顾了精准率和召回率两个主要的指标，其评价更具有实际意义，因此使用 F1-score 来对模型效果进行效果评定。而 F1-score 的评估目标，是目标域中无标记样本中的入侵检测样本是否能够得到有效判别。

- **对照实验组设定**

实验同时设置了对照组，对于迁移学习而言，对照组实验需要从以下几个方面设置：

1. 与直接使用目标域的少量样本训练的模型进行对比；
2. 与不使用迁移学习的方法（使用源域数据训练模型并将之直接应用在目标域数据上）进行对比；
3. 为了研究集成策略是否有效，对照组中也应该包含了训练无集成加权迁移学习模型的实验。

- **互信息计算过程**

在基于集成迁移学习算法的模型中，按照如下方式计算法互信息：计算源域有标记样本的数据分布描述和目标域有标记样本的数据分布描述（数据分布描述，包含均值和标准差，实验过程中使用 python 中 Pandas 库里的 describe()函数生成）之间的互信息值。

- **实验使用的基本分类器参数设置**

本实验中主要使用了两类监督式学习算法：决策树算法和感知算法。其中决策树被使用在了非迁移或非集成的对照组实验中，同时也被使用在了集成迁移学习模型的分类特征集和数值特征集上，而感知机由于较好适用于数值分布在[0,1]之间的特征，因此被使用在集成迁移模型的归一化特征集的迁移模型中。决策树和感知机的基本参数设置如表 3：

表 3 实验使用的模型以及其基本参数

Tab. 3 parameter settings for the experiments

| | 参数名称 | 参数值 |
|---|---|---|
| 决策树参数 | criterion | gini |
| | min_samples_split | 2 |
| 感知机参数 | alpha | 0.0001 |
| | eta | 1.0 |

实验记录了在使用不同量目标域数据情况下，各个特征集在不同域上的互信息值的变化情况，当目标域的有标记样本量比例从 0.1%逐渐扩量到 2.5%的时候，不同特征集下的源域和目标域互信息值如下图 4 所示：

其中，mu_nurm 描述了数值型特征在源域和目标域之间的互信息值，mu_cate 表示类别型特征在源域和目标域之间数据分布的互信息值，mu_norm 表示已经归一化的特征在源域和目标域之间数据分布的互信息值。

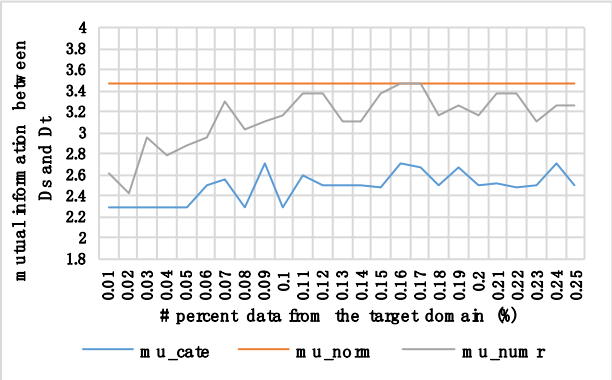


图 4 不同量有标记目标域样本下每个特征集在不同域下的互信息值

Figure 4 mutual information between the source domain and the target domain according to the increcement of the ratios of labeled target domain

由图 4 可知，分类特征和数值特征随着目标域有标记样本量的增加，其源域和目标域之间的互信息值也产生了一些波动，而同时已经被归一化的数值特征由于原始的均值和标准差较为稳定，因此波动较小，以上数据曲线也说明了每一次源域和目标域在某一特征集上的可迁移的知识权重实际是存在差异的。

- **实验结果**

实验统计了随着目标域有标记样本在目标域样本中所占比率逐渐增多的情况下，使用集成迁移模型和使用非集成、非迁移模型的实验结果，节选的部分结果（目标域未标记样本所占比率为 0.1%~1.3%）如表 4 所示。更丰富的实验结果绘制如图 5 和图 6 所示，其中，图 5 描述了集成迁移模型和非集成迁移模型的实验对照，而图 6 显示了集成迁移模型和简单迁移模型的实验对照。

表 4 F1 scores in the four experiments according to the increasement of the ratio of the data from the target domain

Tab. 4 集成迁移模型和非迁移、非集成方法随着目标域样本的量增加时四组模型的预测性能（F1-score）

| 实验策略 | 目标域样本采样比例（%） | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 | 1.1 | 1.2 | 1.3 |
| 使用决策树训练目标域少量有标记样本，并直接预测目标域无标记样本 | 0.78 | 0.79 | 0.79 | 0.81 | 0.81 | 0.83 | 0.83 | 0.83 | 0.84 | 0.84 | 0.86 | 0.86 | 0.87 |
| 使用决策树训练非集成迁移模型，预测目标域无标记样本 | 0.79 | 0.83 | 0.84 | 0.86 | 0.87 | 0.88 | 0.88 | 0.89 | 0.90 | 0.89 | 0.91 | 0.92 | 0.92 |
| 使用决策树在源域训练并直接预测目标域无标记样本 | 0.75 | 0.81 | 0.85 | 0.86 | 0.87 | 0.87 | 0.87 | 0.89 | 0.90 | 0.89 | 0.91 | 0.91 | 0.91 |
| 对分类特征和数值特征使用决策树，对已 | **0.84** | **0.86** | **0.89** | **0.90** | **0.90** | **0.91** | **0.91** | **0.92** | **0.92** | **0.92** | **0.93** | **0.93** | **0.93** |

经归一化的特征使用多层感知机，训练集成迁移模型预测目标域无标记样本

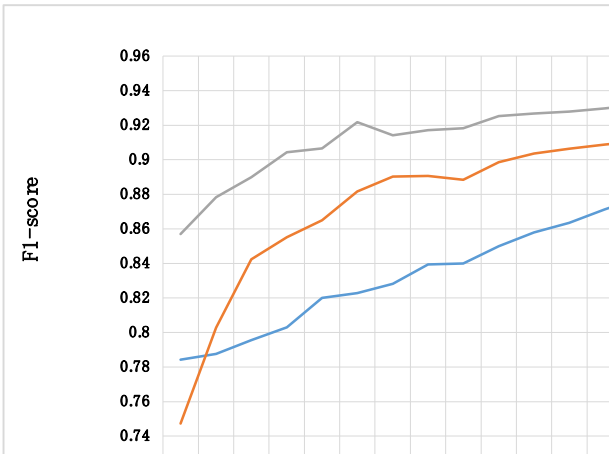


图 5 集成迁移模型和非迁移模型的实验对照

Figure 5 Different performances between ensemble transfer learning method and non-transfer learning methods

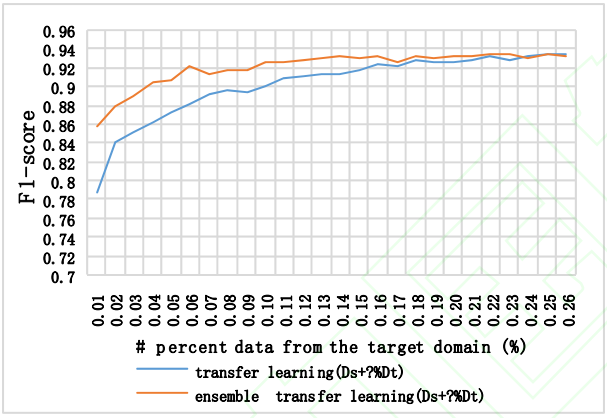


图 6 集成迁移模型和单迁移模型的实验对照

Figure 6 Different performances between ensemble transfer learning method and non-ensemble transfer learning method

- **实验分析**

结合表 4，分析图 5 和图 6 的实验结果，可以得到两组实验的一些基本结论：

由图 5 中的 exp1 可知：仅仅使用源域的数据混合目标域中的少量有标签数据训练分类模型，在目标域的无标记样本上的预测结果并不佳，当目标域中的有标记样本在 0.02%~0.25%区间段递增时，这种策略甚至比不上直接使用目标域的少量有标记样本直接训练的模型（图 5 中 exp2）的效果好。

在同样的实验条件下，使用了集成迁移学习模型的预测结果，是优于以上两种不使用迁移策略的方法的：在目标域样本极少（0.1%）时，集成迁移模型的预测效果在一开始较高，这说明极大地利用了目标域的有限样本筛选了更多对目标域有利的源域样本，并据此辅助训练目标域模型，这也说明有选择地迁移源域知识对训练目标域的模型是有利的。

另外，对比图 6 中的两组曲线可知，对多个特征集分别训练迁移模型并使用互信息加权集成多个模型的效果，是好于不使用集成策略的迁移学习模型的。虽然在目标域有限样本量更多的情况下，不使用集成策略的迁移模型会慢慢追平使用集成策略的迁移模型的 F1-score 值，但是在样本量更少的情况下，集成迁移模型可以具有更强的提早发现入侵检测异常的能力，而这一点非常适用于真实环境下的网络入侵检测环境。

因此，结合以上的实验结论可知，在基于实际场景下跨领域的入侵检测分类问题中，完全混合源域和目标域的数据训练模型以及直接使用目标域少量数据训练模型，不会对目标任务有利，而基于互信息加权的集成迁移模型对于目标域的模型训练有利。

## 4 总结

通过将迁移学习技术和集成学习的思想应用在入侵检测领域，对源域和目标域多组不同的特征集，使用简单的迁移策略，训练较好的独立的迁移模型；然后使用互信息衡量源域和目标域在该特征集下的数据分布差异并以之对多个迁移模型进行集成加权，得到最终的集成迁移模型。通过在 NSL-KDD 标准的入侵检测数据集中的实验得知，该集成迁移模型的效果好于不使用迁移模型的效果，同时也好于不使用集成策略的迁移模型的效果。

迁移的本质是挖掘源域中的可用知识来辅助目标域决策，虽然使用集成模型的方法已经一定程度上优化了对目标域的分类模型的学习。但是，对于分类特征而言，其仍然包含了一些没有得到有效解析的自然语言相关的信息，未来如何引入 NLP（Natural Language Processing）相关的知识来对这些特征进行迁移也是解决入侵检测问题的新视角之一。

**参考文献**

[1] Akhil J , Sultana A . Intelligent Network Intrusion Detection System using Data Mining Techniques[C]// International Conference on Applied & Theoretical Computing & Communication Technology. IEEE, 2017.

[2] Ahmadi R, Macredie R D, Tucker A. Intrusion Detection Using Transfer Learning in Machine Learning Classifiers Between Non-cloud and Cloud Datasets[C]//International Conference on Intelligent Data Engineering and Automated Learning. Springer, Cham, 2018: 556-566.

[3] Aljawarneh S, Aldwairi M, Yassein M B. Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model[J]. Journal of Computational Science, 2018, 25: 152-160.

[4] 陈虹, 万广雪, 肖振久. 基于优化数据处理的深度信念网络模型的入侵检测方法[J]. 计算机应用, 2017(6).

[5] Liu J, He J, Zhang W, et al. ANID-SEoKELM: Adaptive network intrusion detection based on selective ensemble of kernel ELMs with random features[J]. Knowledge-Based Systems, 2019, 177: 104-116.
[6] Benmessahel I, Xie K, Chellal M, et al. A new evolutionary neural networks based on intrusion detection systems using locust swarm optimization[J]. Evolutionary Intelligence, 2019: 1-16.
[7] 高妮, 贺毅岳, 高岭. 海量数据环境下用于入侵检测的深度学习方法[J]. 计算机应用研究, 2018, 35(4): 1197-1200
[8] 汪世义, 陶亮, 王华彬. 几种机器学习方法在 IDS 中的性能比较[J]. 计算机仿真, 2010, 27(8):92-94.
[9] 刘冬兰, 马雷, 刘新, 等. 基于深度学习的电力大数据融合与异常检测方法[J]. 计算机应用与软件, 2018, 35(4): 61-64.
[10] Xiao J , Wang R , Teng G , et al. A Transfer Learning Based Classifier Ensemble Model for Customer Credit Scoring[C]// 2014 Seventh International Joint Conference on Computational Sciences and Optimization (CSO). IEEE, 2014.
[11] Sun Q , Amin M , Yan B , et al. Transfer Learning for Bilingual Content Classification[J]. 2015.
[12] Shao L, Zhu F, Li X. Transfer learning for visual categorization: A survey[J]. IEEE transactions on neural networks and learning systems, 2014, 26(5): 1019-1034.
[13] Tang J, Zhao Z, Bei J, et al. The application of transfer learning on e-commerce recommender systems[C]//2013 10th Web Information System and Application Conference. IEEE, 2013: 479-482.
[14] Huynh B Q, Li H, Giger M L. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks[J]. Journal of Medical Imaging, 2016, 3(3): 034501.
[15] Zhao H, Liu Q, Yang Y. Transfer learning with ensemble of multiple feature representations[C]//2018 IEEE 16th International Conference on Software Engineering Research, Management and Applications (SERA). IEEE, 2018: 54-61.
[16] Dai W , Yang Q , Xue G R , et al. [ACM Press the 24th international conference - Corvalis, Oregon (2007.06.20-2007.06.24)] Proceedings of the 24th international conference on Machine learning - ICML \"07 - Boosting for transfer learning[C]// International Conference on Machine Learning. ACM, 2007:193-200.
[17] Javaid A, Niyaz Q, Sun W, et al. A deep learning approach for network intrusion detection system[C]//Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016: 21-26.
[18] Ambusaidi M A, He X, Nanda P, et al. Building an intrusion detection system using a filter-based feature selection algorithm[J]. IEEE transactions on computers, 2016, 65(10): 2986-2998.
[19] Vinayakumar R, Soman K P, Poornachandran P. Applying convolutional neural network for network intrusion detection[C] //In 2017 International Conference on Advanced Computing, Communications and Informatics, Piscataway, New Jersey, USA:IEEE, 2017: 1222-1228.
[20] NSL-KDD dataset [EB/OL]. [2018-07-20] https://www.unb.ca/cic/datasets/nsl.html
[21] Kabir E, Hu J, Wang H, et al. A novel statistical technique for intrusion detection systems[J]. Future Generation Computer Systems, 2018, 79: 303-318.

胡健（1992-），男，云南文山人，工程师，硕士 ，主要研究方向：信息安全、机器学习；苏永东(1967—)，女，北京市人，高级工程师，主要研究方向：信息安全；黄文载(1963—)，男，云南昆明人，高级工程师，主要研究方向：电力系统自动化；肖鹏(1988—)，男，云南昆明人，工程师，主要研究方向：网络空间安全；刘玉婷(1987—)，女，云南昭通人，工程师，硕士研究生，主要研究方向：信息安全；杨本富(1982—)，男，云南保山人，工程师，主要研究方向：软件工程、信息安全。

HU Jian, born in 1992, M. S., His research interests include information security, machine learning;ZHOU Jin, born in 1973, His research interests include automation of electric power system; SU Yongdong, born in 1967, Her research interests include network information security; HUANG Wenzai, born in 1963, His research interests include automation of electric power system; XIAO Peng, born in 1988, His research interests include Cyberspace Security;LIU Yuting, born in 1987, M. S., Her research interests include information security; YANG Benfu, born in 1982, His research interests include software engineering, information security.