



一种基于不均衡数据的网络入侵流量分类方法

关晓惠¹, 钱亚冠²

(1. 浙江水利水电学院 杭州 310018; 2. 浙江科技学院理学院 杭州 310023)

摘要:在网络入侵流量检测中,普遍存在不同攻击类型的流量分布不均现象,导致少数攻击流量类识别率较低。为解决此类问题,基于不同特征空间的分类器流水线组合方法将多分类问题转化为不同特征空间上的两分类问题,有效地实现少数类重抽样和特征空间的优化,避免了少数类受多数类特征的干扰。实验表明,此方法可以有效地提高攻击流量中少数类的分类精度和召回率。

关键词:攻击流量;类不均衡;分类器流水线组合

doi: 10.11959/j.issn.1000-0801.2015085

A Network Traffic Classification Method for Class-Imbalanced Data

Guan Xiaohui¹, Qian Yaguan²

(1. Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China;

2. College of Science, Zhejiang University of Science and Technology, Hangzhou 310023, China)

Abstract: It is very common that flow distribution of class is not uniform in attack traffic. It will lead to a low classification accuracy in network intrusion detection. For overcoming this class imbalance phenomenon, a pipelining ensemble approach in different feature spaces was proposed, which translates multi-class classification to two-class classification. Based on the pipelining ensemble, it could be further conduct oversampling and customized feature selection for minority class, which may avoid the disturbance from majority class. The experiment result shows that the proposed approach can efficiently improve the accuracy of minority class of attack traffic.

Key words: attack traffic, class imbalance, classifiers pipelining ensemble

1 引言

目前,各种机器学习方法已广泛应用于网络攻击流量的识别中,并获得了不错的分类效果。但是这些方法大都

将提高整体准确率(overall accuracy rate)或者多数类(majority class)的分类准确率作为方法有效性的评价标准。以往的研究大都忽略了一个重要的事实,即互联网中的各种攻击流量分布是极不均匀的,称为类不均衡现象。

收稿日期:2014-12-04;修回日期:2015-04-07

基金项目:浙江省水利厅科技计划基金资助项目(No.RC1452),浙江省网络媒体云处理与分析工程技术中心开放课题基金资助项目(No.2012E10023-14),2014年度高校国内访问学者专业发展基金资助项目(No.FX2014092)

Foundation Items: Water Resources Department Foundation of Zhejiang Province(No.RC1452), The Open Project of Cloud Processing and Analysis Center of Network Media of Zhejiang Province(No.2012E10023-14), The Professional Development Project of 2014 College Visiting Scholar(No.FX2014092)

论文引用格式:关晓惠,钱亚冠.一种基于不均衡数据的网络入侵流量分类方法.电信科学,2015085

Guan X H, Qian Y G. A network traffic classification method for class-imbalanced data. Telecommunications Science, 2015085

表 1 给出了 KDD cup99^[1]数据集中不同攻击类型的流量比例,可以明显发现 DoS 攻击在数据流比例上占有绝对优势,而 U2R 等所占比例极小。这种不均匀的比例分布,使得以数据流为统计对象的分类方法产生不符合实际情况的准确率。例如表 1 中,即使将所有的其他攻击类型错误地预测到 DoS,也可以获得 98.33% 的整体准确率。这样的“高”准确率显然没有实际意义,因为其他所有攻击类型的分类准确率均为 0。NaïveBayes、C4.5、 k -NN、SVM 4 种机器学习方法对 KDD cup99 网络入侵数据分类准确性的比较结果如图 1 和图 2 所示。可以看出,不同比例的攻击类型,其分类准确率的差异非常显著。这种目标类比例严重失衡而导致少数类识别误差增大的问题被称为类不平衡(class imbalance)问题。如何解决类不平衡问题已经成为机器学习方法成功应用于攻击流量分类的一个关键性问题。本文提出了一种基于不同特征空间的分类器流水线组合方法,将多分类问题转化为不同特征空间上的两分类问题,有效解决了分类样本的不均衡问题,从而提高了分类的整体准确率。

表 1 KDD cup99 数据集中的攻击流量分布比例

攻击类型	数据流数量/个	比例
DoS	39 092	98.33%
U2R	13	0.03%
R2L	213	0.54%
probe	437	1.10%

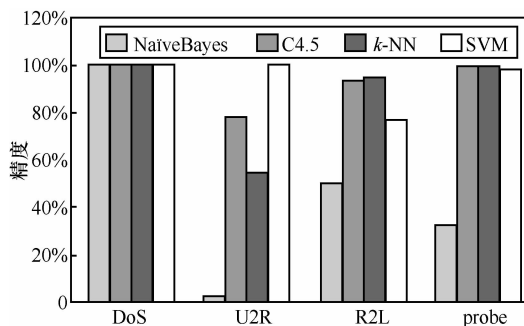


图 1 4 种机器学习方法的精度比较

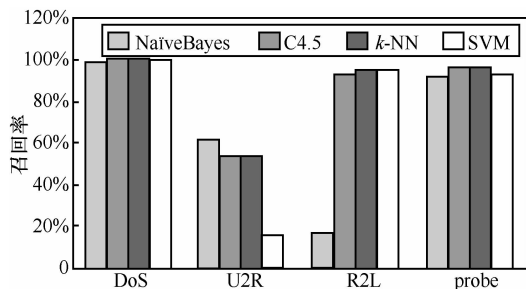


图 2 4 种机器学习方法的召回率比较

2 相关工作

类不平衡问题在机器学习领域已经引起了学者们的广泛关注,目前主要的解决方法是对数据重抽样(resampling)、特征选择与分类器组合(classifiers ensemble)。

2.1 重抽样

对原始训练数据集进行重抽样,形成新的数据集,其目的是为了让不同分类的不均衡分布趋向于均衡分布。Weiss 和 Provost^[2]注意到,原始的样本分布不一定是给定分类器的最优分布,而通过重抽样方法可以使样本的分布接近最优。

重抽样技术主要包括对多数类的欠抽样(under-sampling)^[3,4]、对少数类的过抽样(over-sampling)^[5,6]及联合欠抽样与过抽样^[7]。简单的随机欠抽样与过抽样都具有自身的不足:对多数类的欠抽样会导致一些对分类器有价值的信息丢失;而对少数类的简单重复抽样对于提高分类性能并无太大的帮助^[8],甚至会导致分类器对数据的过度拟合(overfit)^[9]。为此,Chawla 等人^[6]提出了新的过抽样技术——SMOTE 算法,其基本思想是通过人工合成新的少数类样本来减轻类别的不均衡,解决传统过抽样技术因决策域变小而引起的过拟合现象。

2.2 特征选择

因为类的不均衡问题通常伴随着数据的高维特征问题,因此特征选择是必须的步骤。Putten 等人^[10]认为,良好的特征选择方法比优化分类算法本身对于提高分类性能具有更高的价值。特征选择方法已经被研究了很长时间,在近期开始被研究人员用于解决类不平衡问题^[11,12]。

Zhang 等人^[13]专门研究了特征选择方法应用于网络流量中的类不平衡问题,提出了一种新的特征选择评价指标,称为加权系统不确定性(weighted symmetrical uncertainty, WSU)。将该评价指标与 AUC(area under roc curve)结合,设计了一种新的 wrapper 型的特征选择算法。在 Moore 提供的流量数据上进行了验证,结果表明,该算法可使分类器获得的流准确率大于 94%,字节准确率大于 80%。

2.3 分类器组合

组合学习是当前机器学习中主要的热点研究方向之一,它的根本思路是对同一问题使用一系列分类器进行学习,并使用一定的策略把各个不同的分类结果进行整合,从而获得比单个分类器更好的分类效果^[15-18]。Galar 等人^[14]系统地综述了最近的研究成果,研究了这些方法之间的关系,并指出组合方法必须联合重抽样等预处理方法才能更



好地解决类不平衡问题。这些分类器组合方法都是在相同特征空间的基础上进行的。

3 分类器流水线组合方法

研究发现,同一特征对不同流量类的区分能力是不同的^[20],如数据分组大小的均值可以较好地区分 SSH 和 P2P 应用,但却不能很好地区分 FTP 和 P2P 应用。在 KDD cup99 的攻击流量中同样存在类似的性质,因此本文提出多个两分类器的流水线组合方法。与以往并行组合分类器的方式不同,本文的方法将多个两分类器进行串行组合,每个两分类器只负责区分某一个类,无法识别的流量传递到下一个分类器完成。由于每个分类器只负责识别某一种攻击流量,尤其是少数类的流量,可以对它们的特征空间进行专门的优化,而不是将多个分类器统一到相同的特征空间。

3.1 数据预处理

本文将区分 DoS、U2R、R2L 与 probe 这 4 种攻击流量的多分类问题分解为 {DoS, others}、{U2R, others}、{R2L, others}、{probe, others} 这 4 个两分类问题。因此,需要对原始的 4 类样本数据进行预处理。

考虑到攻击流量类型的不均衡性,首先对原始训练数据集 D 中少数类 U2R、R2L 等进行过抽样,使其达到一定的比例。然后将所有非 DoS 的流量类标注为 others 流量类,形成只有 DoS 与 others 这两类流量的数据集 D_{DoS} 。接下来将原始训练集中的所有 DoS 流量类样本实例过滤掉,将余下的非 probe 流量类全都标记为 others 流量类,形成只有 probe 与 others 这两类流量的数据集 D_{probe} 。以同样的方式,从原始训练集中过滤掉所有的 DoS、probe 流量实例后,将剩下的非 R2L 流量标记为 others 流量,形成新的训练集合 D_{R2L} 。以此类推,最后获得训练集合 D_{U2R} 。 D_{DoS} 、 D_{probe} 、 D_{R2L} 与 D_{U2R} , 这 4 个训练集合分别用于训练两分类器 $M_{DoS/others}$ 、 $M_{probe/others}$ 、 $M_{R2L/others}$ 与 $M_{U2R/others}$ 。

3.2 特征选取

常规的分类方法在区分 DoS、U2R、R2L 与 probe 这 4 种攻击流量时,先通过某种评价函数获得某个特征空间 $\mathcal{F}=\{a_1, a_2, \dots, a_k\}$, 在这个特征空间下学习获得某种分类模型。但用于区分 4 种攻击流量的特征空间 $\mathcal{F}$ 不一定是区分 U2R 的最佳特征空间,可能 $\{a_1, a_2\}$ 是它的最佳特征空间,余下的 $\{a_3, \dots, a_k\}$ 反而可能成为干扰特征。因此,本文为每种攻击流量提取各自的最佳特征空间,将传统上采用单一特征空间的分类策略转变为多特征空间下的分类策略。

训练样本预处理完成后,需要为 {DoS, others}、{U2R, others}、{R2L, others}、{probe, others} 这 4 种二分类问题分别提取相应的特征空间 $\mathcal{F}_{DoS}$ 、 $\mathcal{F}_{U2R}$ 、 $\mathcal{F}_{R2L}$ 、 $\mathcal{F}_{probe}$ 。本文使用 Relief 算法,对各个特征赋予不同的权重,仅选择权重大于某个阈值的特征参与分类器的训练。算法思想是从训练样本集中随机选择样本 R , 然后从和 R 同类的样本中寻找最近邻样本 H , 称为 *NearHit*, 最后从和 R 不同类的样本中寻找最近邻样本 M , 称为 *NearMiss*。如果 R 和 *NearHit* 在某个特征上的距离小于 R 和 *NearMiss* 上的距离,说明该特征对区分两类的最近邻是有益的,则增加该特征的权重;反之,如果 R 和 *NearHit* 在某个特征的距离大于 R 和 *NearMiss* 上的距离,说明该特征对区分两类的最近邻起负面作用,则降低该特征的权重。以上过程重复 m 次(m 为选择的样本个数),最后得到各特征的平均权重。特征的权重越大,表示该特征的分类能力越强;反之,表示该特征分类能力越弱。

3.3 分类器的训练

在预处理后的 4 个训练样本集 D_{DoS} 、 D_{probe} 、 D_{R2L} 、 D_{U2R} 用上述的方法进行特征选取,形成 4 个独立的特征空间 $\mathcal{F}_{DoS}$ 、 $\mathcal{F}_{probe}$ 、 $\mathcal{F}_{R2L}$ 、 $\mathcal{F}_{U2R}$, 分别作为 4 个两分类器的特征空间。每个分类器模型分别在各自的特征空间上进行训练。数据预处理、特征选取和分类器训练流程如图 3 所示。

在特定分类器训练之前,首先要为该分类器选择最优的特征空间,具体过程如算法 1 所示。算法的时间复杂度为 $O(mn)$ 。

算法 1 Feature_Selection。

输入:已标注过的流量数据 T , 全部特征的集合 A , 阈值 δ , 样本实例的个数 m , 特征集的个数 n 。

输出:特征子集 A' 。

设置每个特征权值 $w(j) \leftarrow 0$, A' 为空集;

for $i \leftarrow 1$ to m

从 T 中随机选取一个样本实例 R ;

从 R 的同类样本集中寻找最近邻样本实例 H ;

从 R 的不同类样本集中找到最近邻的样本实例 M ;

for $i \leftarrow 1$ to m

$w(j) \leftarrow w(j) - \text{diff}(j, R, H) / m + \text{diff}(j, R, M) / m$; // 更新权值

新权值

end for

end for

for $j \leftarrow 1$ to n

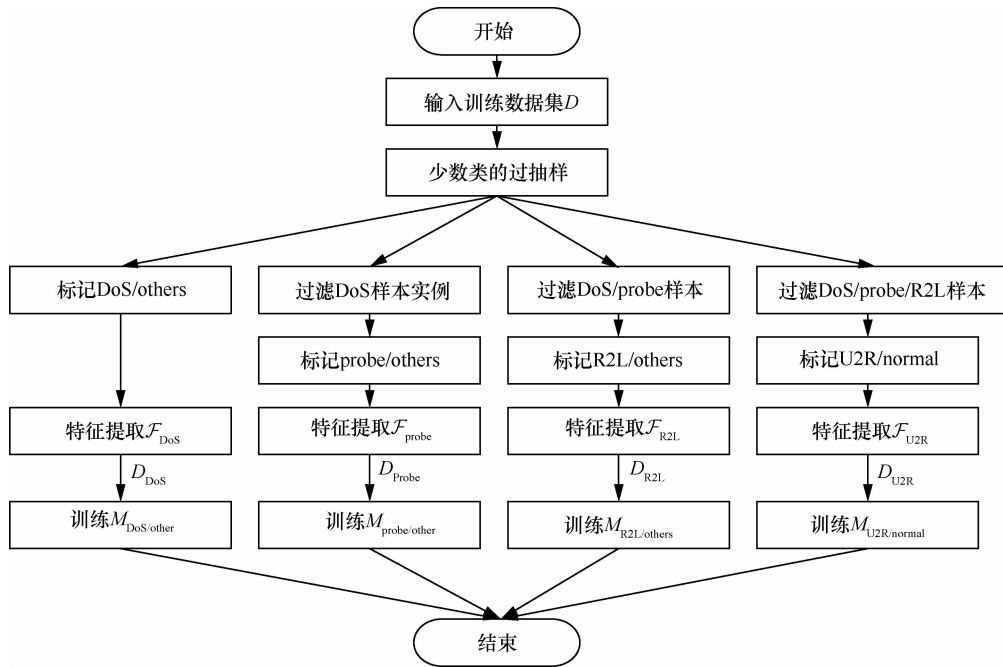


图3 KDD cup99 数据集预处理过程及分类器训练流程

if $w(j) > \delta$, 把第 j 个特征添加到 A' 中;

end for

返回 A'

分类器的训练过程如算法 2 所示。在训练阶段的时间复杂性除了受特征选择影响之外, 还和两分类器的选择有关。为了提高分类效率和减少训练时间, 选择 Naïve Bayes、C4.5、 k -NN、SVM 等经典算法作为两分类器算法。以 Naïve Bayes 为例, 整个训练阶段的时间复杂度为 $O(mnk)$, 其中 m 为训练实例的个数, n 为对应特征空间的特征个数, k 为类别数。

算法 2 Training_Binary_Classifiers.

输入: D 是已标记过的流量数据, F 是全部特征的集合。

输出: 两分类器 $\{M_i | i=1, \dots, k\}$ 。

for $i \leftarrow 1$ to k // 假设有 k 个攻击流量类

D 中除类标签为 C_i 的实例外, 其余训练实例全都标记为 others, 形成训练集合 D_i ;

$F'_i \leftarrow \text{Feature-Selection}(T_i, F)$; // 得到特征子集 F'_i

根据训练集 D_i , 特征子集 F'_i 学习得到两分类器 M_i ;

for $j \leftarrow 1$ to i

$D \leftarrow$ 从 D 中过滤到所有类标签为 C_j 的实例; // 准备下一个分类器的训练数据

end for

end for

return $\{M_i | i=1, \dots, k\}$

如图 3 所示, 共训练了 4 个分类器, 每个分类器只区分两类流量, 如 $M_{\text{DoS/others}}$ 区分 DoS 和非 DoS 流量。它们存在的主要区别是: others 流量类的流量组成不同; 特征空间不同。在整个分类器流水线组合过程中, 越靠前的分类器, 其训练集中 others 流量类包含的流量种类越多, 如 D_{DoS} 中的 others 流量至少包含 probe 流量、R2L 流量、U2R 流量等。但对于分类器 $M_{\text{DoS/others}}$ 而言, 这些流量均被标记为 others 流量, 即非 DoS 流量。对于 $M_{\text{probe/others}}$ 而言, 因为前面的 $M_{\text{DoS/others}}$ 已经过滤掉了 DoS 流量, 所以它并不需要再负责识别 DoS 流量。因此, $M_{\text{probe/others}}$ 的训练集合 D_{probe} 中不需要包含 DoS 流量实例, 即 others 流量只包含 R2L 流量、U2R 流量及正常流量。这样做的好处是越到后面的分类器, 其 others 流量构成成分越少, 受其他类的干扰越小, 分类就越容易。同时, 因为训练集的构成发生了变化, 所以需要对这些训练集重新进行特征选择, 最后形成了每个分类自己的特征空间。

3.4 分类过程

本文采用流水线串行化的方式将训练好的单一分类器模型 $M_{\text{DoS/others}}$ 、 $M_{\text{probe/others}}$ 、 $M_{\text{R2L/others}}$ 和 $M_{\text{U2R/others}}$ 进行流水线组合, 如图 4 所示, 前一个分类器的输出将作为后一个分类器的输入。4 个分类器的组合次序是按照其分类难度排列的, 即 $M_{\text{DoS/others}} \xrightarrow{\text{others}} M_{\text{probe/others}} \xrightarrow{\text{others}} M_{\text{R2L/others}} \xrightarrow{\text{others}} M_{\text{U2R/others}}$, 越是难以识别的类别就安排在流水线越靠前的部位。分类时, 将一个新样本输入组合分类器, 分类器 $M_{\text{DoS/others}}$ 首先对输入的样本

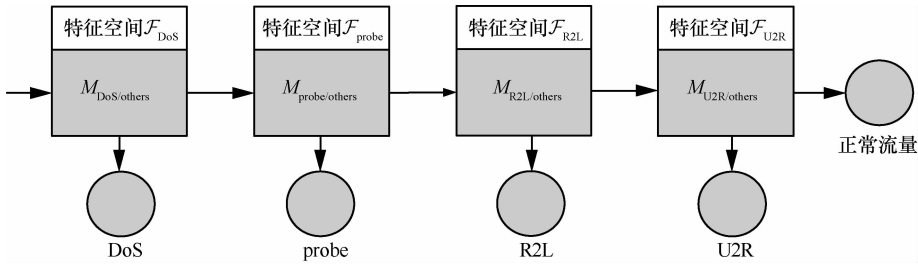


图4 KDD cup99数据集流水线组合分类器的分类流程

进行判别,将其预测为 DoS 流量或 others 流量。如果预测为 others 流量,则该样本继续作为下一个分类器 $M_{probe/others}$ 的输入,如此下去,直至其被作为 R2L 流量、U2R 流量或正常流量之一输出,具体过程如算法 3 所示。

算法 3 Cascade_Classify。

输入:测试数据集 T ,类别 $\{C_1, C_2, \dots, C_t\}$,训练好的两分类器 $\{M_1, M_2, \dots, M_k\}$,对应两分类器 $\{M_1, M_2, \dots, M_k\}$ 的特征空间 $\{F_1, F_2, \dots, F_{t-1}\}$ 。

输出:类别 $\{C_i\}$ 。

```
while ( $t \leftarrow \text{getNext}(T)$ )  $\neq$  null //从测试集中获取数据
    for  $i \leftarrow 1$  to  $k$ 
         $t' \leftarrow$  根据特征空间  $S_i$  上提取测试样本的特征;
         $C \leftarrow M_i(t')$ ; //分类器  $M_i$  对  $t'$  的类别进行预测;
        if  $C = C_i$  then 输出  $C$ ;
        else 转至(2);
    end if
end for
输出 normal; //预测为正常流量
end while
```

分类过程的时间耗费与两分类器选择有关。以 Naïve Bayes 为例,最好的情况是测试样本为 DoS 流量,只经过第一个分类器就可以输出,时间复杂度为 $O(mnk)$,最坏的情况是测试样本需经过 4 个分类器才能输出结果,时间复杂度为 $O(4mnk)$ 。

4 实验分析

本文提出的流水线组合方法可以应用任何已有的分类算法作为两分类器。使用 Naïve Bayes、C4.5、 k -NN、SVM 这 4 种分类算法作为基本的两分类器。为了便于比较分类器组合后的性能,只将同一类算法的分类器进行组合,例如将 4 个 Naïve Bayes 两分类器进行串行组合。图 5~图 8 给出了上述 4 种分类算法构建成的 4 个流水线组合分类器与单一多分类器的性能比较,分类性能的评估策略采用 10 折交叉验证的方法比较精度与召回率。图 5 对比了单一特征空间下的 Naïve Bayes 多分类器和多特征空间下由 Naïve Bayes 作为基本分类器构建的流水线组合分类器的分类精度与召回率。可以发现 U2R 流量作为少数类,其精度和召回率得到了大幅提高,精度的提高幅度超过了 60%,召回率提高幅度也在 20% 左右。

同样,从图 6、图 7、图 8 中可以明显观察到,这些基本分类器在采用流水线方式组合后,其准确率得到了有效提升。而横向比较 4 种组合分类器的性能可以发现,C4.5 和 k -NN 实现流水线组合后可使少数类 U2R、R2L 等流量类的精度和召回率达到 90% 以上,优于 Naïve Bayes 和 SVM 流水线组合后分类识别性能。由此可以看出,流水线组合后可以实现单个分类器在少数类上无法达到的准确率。

本分类算法是以流水线的形式组合已有的经典分类器,充分利用不同分类器的优势。每个两分类器可以相同

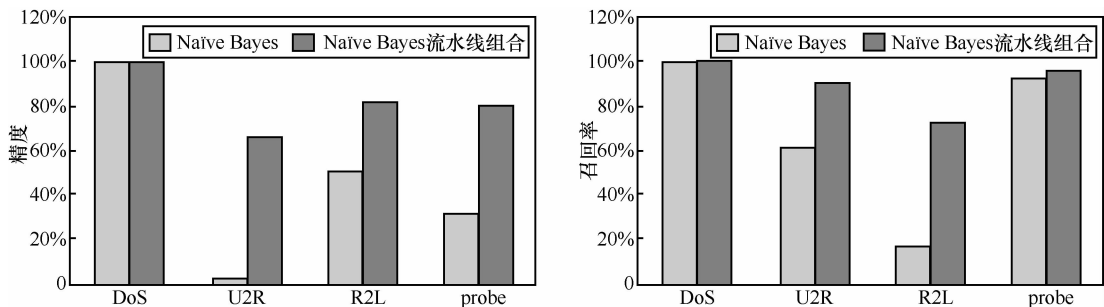


图5 NaïveBayes 分类器流水线组合后的精度与召回率

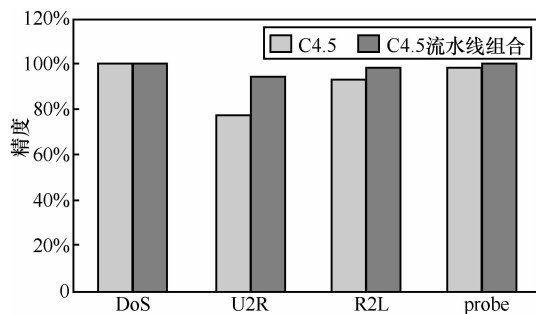


图6 C4.5分类器流水线组合后的精度与召回率

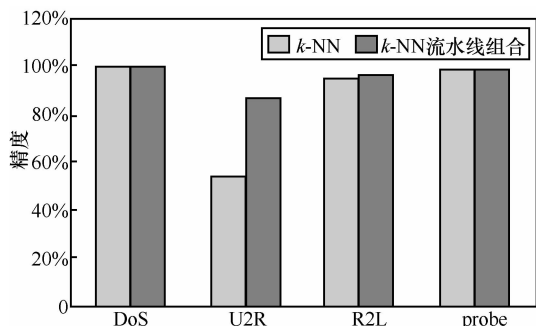
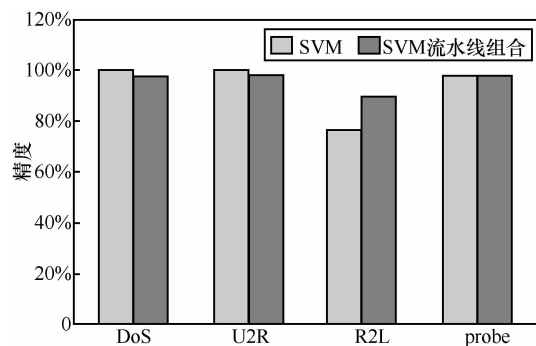
图7 k -NN分类器流水线组合后的精度与召回率

图8 SVM分类器流水线组合后的精度与召回率

也可以不同。实验证明,该方法对样本比较少的类别的分类正确率提高得比较多。算法在训练阶段和分类阶段都有较低的时间复杂度。

5 结束语

机器学习方法应用于攻击流量的分类识别被证明是一种很有前途的方法。但目前该方法在应用过程中缺乏对类不平衡问题的研究,而类不平衡问题又是网络攻击流量中普遍存在的现象。本文提出基于多特征空间的流水线组合方法,将多分类问题简化为两分类问题,实现对少数类特征空间的优化,有利于少数类的识别。这种分类框架可以与任何一种基本的分类算法结合,具有很强的灵活性。实验结果表明,这种流水线组合方法可以有效提升对少数

类的识别准确率。

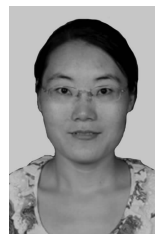
参考文献

- 1 Hettich S, Bay S D. KDD cup 1999. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>, 2007
- 2 Weiss G, Provost F. Learning when training data are costly: the effect of class distribution on tree induction. *Artificial Intelligence Research*, 2003(19): 315-354
- 3 Kubat M, Matwin S. Addressing the curse of imbalanced data sets: one sided sampling. *Proceedings of the 14th International Conference on Machine Learning*, Nashville, USA, 1997
- 4 Chen X, Gerlach B, Casasent D. Pruning support vectors for imbalanced data classification. *Proceedings of IEEE International Joint Conference on Neural Networks*, Montreal, Quebec, Canada, 2005



- 5 Kubat M, Matwin S. Learning when negative examples abound. Proceedings of the 9th European Conference on Machine Learning, Prague, Czech Republic, 1997
- 6 Chawla N, Bowyer K, Hall L, *et al*. SMOTE: synthetic minority over-sampling technique. Artificial Intelligence Research, 2002(16): 321~357
- 7 Estabrooks A, Jo T, Japkowicz N. A multiple resampling method for learning from imbalanced data sets. Computational Intelligence, 2004, 20(1): 18~36
- 8 Ling C, Li C. Data mining for direct marketing problems and solutions. Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, New York, USA, 1998
- 9 Chawla N, Japkowicz N, Kotcz A. Editorial: special issue on learning from imbalanced data sets. ACM SIGKDD Explorations Newsletter, 2004, 6(1): 1~6
- 10 Putten P V D, Someren M V. A bias-variance analysis of a real world learning problem: the CoIL challenge 2000. Machine Learning, 2004, 57(1-2): 177~195
- 11 Forman G. An extensive empirical study of feature selection metrics for text classification. Machine Learning Research, 2003(3): 1289~1305
- 12 Zheng Z, Wu X, Srihari R. Feature selection for text categorization on imbalanced data. ACM SIGKDD Explorations Newsletter, 2004(6): 80~89
- 13 Zhang H, Lu G, Qassrawi M T, *et al*. Feature selection for optimizing traffic classification. Computer Communications, 2012, 35(12): 1457~1471
- 14 Galar M, Fernández A, Barrenechea E, *et al*. A review on ensembles for the class imbalance problem: bagging, boosting, and hybrid-based approaches. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2012, 42(4): 463~484
- 15 Seiffert C, Khoshgoftaar T, Van Hulse J, *et al*. Rusboost: a hybrid approach to alleviating class imbalance. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 2010, 40(1): 185~197
- 16 Liu X Y, Wu J, Zhou Z H. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2009, 39(2): 539~550
- 17 Sun Y, Kamel M S, Wong A K, *et al*. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition, 2007, 40(12): 3358~3378
- 18 Van Hulse J, Khoshgoftaar T, Napolitano A. An empirical comparison of repetitive undersampling techniques. Proceedings of IEEE International Conference on Information Reuse & Integration, Las Vegas, NN, USA, 2009: 29~34
- 19 Breiman L. Bagging predictors. Machine Learning, 1996 (24): 123~140
- 20 Xie G, Iliofotou M, Keralapura R, *et al*. Subflow: towards practical flow-level traffic classification. Proceedings of IEEE INFOCOM, Orlando, FL, USA, 2012

[作者简介]



关晓惠,女,浙江水利水电学院副教授,主要研究方向为机器学习与数据挖掘。

钱亚冠,男,博士,浙江科技学院理学院副教授,主要研究方向为互联网流量分类、下一代互联网、机器学习与数据挖掘。

2015085-7

(上接 2015146-4)

- Ma Y, Xiao K T, Wang X. The climate characteristics of tarim basin of the sand dust weather. Journal of Beijing University, 2006, 42(6): 784~790
- 5 魏文寿, 高卫东, 史玉光. 新疆地区气候与环境变化对沙尘暴的影响研究. 干旱区地理, 2004, 27(2): 138~141
Wei W S, Gao W D, Shi Y G. Xinjiang regional climate and environment change on the influence of the dust storm study. Journal of Arid Areas in Geography, 2004, 27(2): 138~141
 - 6 Maciel L R, Bertoni H L, Xia H H. Unified approach to prediction of propagation over buildings for all ranges of base station antenna height. IEEE Transactions on Vehicular Technology, 1993, 42(1): 41~45

2015146-5

[作者简介]



陈杰,女,塔里木大学信息工程学院讲师,主要研究方向为移动通信和信号处理。

姚娜[通讯作者],女,塔里木大学信息工程学院副教授,主要研究方向为图像处理、农业信息化的研究。