基于迁移学习和 D-S 理论的网络异常检测*

赵新杰，刘 渊，孙 剑
（江南大学 数字媒体学院，江苏 无锡 214122）

摘 要： 对于分布不同或分布相似的未知类型的网络攻击，目前的异常检测技术往往不能达到预期的效果。针对上述问题，研究了一种基于迁移技术和 D-S 证据理论的网络异常检测方法。首先用迁移学习方法对已知网络攻击进行建模，此模型在构建时考虑了不同分布的异常攻击间的差异；然后用其训练得到的分类器对未知的网络行为进行分析，结合 D-S 证据理论，可以检测出分布不一致的未知攻击类型。实验结果表明，该方法泛化了传统的网络异常检测技术，对未知的网络异常有着较高的检测率。

关键词： 迁移学习；D-S 理论；异常行为分析；数据融合
中图分类号： TP393.08 **文献标志码：** A **文章编号：** 1001-3695(2016)04-1137-04
doi:10.3969/j.issn.1001-3695.2016.04.039

New network anomaly detection using transfer learning and D-S theory

Zhao Xinjie, Liu Yuan, Sun Jian
(*School of Digital Media, Jiangnan University, Wuxi Jiangsu* 214122, *China*)

Abstract: The current approaches of anomaly detection cannot effectively detect unknown network attacks, which follow the same or different distribution. To solve this problem, this paper proposed a new network anomaly detection using transfer learning technique and D-S theory. At first, this paper created a model for known network attacks with transfer learning method, which considered the distinctions in anomaly attacks following different distribution. Secondly, combined with D-S theory, the classifier could pinpoint unknown network attacks as outliers. The results show that the proposed detection approach has a higher detection rate for unknown network anomalies.

Key words: transfer learning; D-S theory; abnormal behavior analysis; data fusion

0 引言

随着互联网的应用越来越广泛，网络安全成为了人们日益关心的话题，异常行为检测则是网络安全中的一个重要组成部分。就现阶段而言，网络数据流依然是探测网络攻击行为的最佳数据源。

迄今为止，对网络异常行为的检测方法主要有概率统计分析方法[1]、数据挖掘方法[2]、神经网络方法[3]、模糊数学理论[4]、人工免疫方法[5]等。概率统计分析方法的优点是有着非常成熟的概率统计的理论作为支撑，但缺点也十分明显，由于用户的行为具有相当的复杂性，如果想要完全匹配用户的行为会变得十分困难，并且在阈值的确定方面也有很大的难度，这两个因素综合起来会导致误报率和漏报率都比较高。数据挖掘方法可以检测出一些未知的攻击，但是由于它需要大量的主机审计技术或应用程序的日志文件作为数据挖掘的基础，实际计算的复杂度较高，误报率也较高，较难运用到准实时环境中。神经网络方法能够自动学习和更新，有着很好的抗干扰能力，但是它也存在诸如网络拓扑结构和各元素的权重都较难确定的缺点。模糊数学理论也是一种比较好的网络分析手段，但是一般在网络异常检测中需结合其他检测方法，通过模糊度和阈值的比较得到最终的结果。人工免疫方法克服了神经网络和遗传算法的局部收敛的缺陷，具有处理复杂问题的能力，但是人工免疫方法的运算量较大，对样本的空间分布要求较高，压缩阈值不易确定。基于上述算法的优缺点，本文将迁移学习的概念和 D-S 理论引入到网络异常行为检测中来。

迁移学习是运用已知的知识对不同但相关的领域问题进行求解的一种新的机器学习方法[6]。由于迁移学习放宽了传统机器学习中的两个基本假设，能够迁移已有的知识来解决目标领域中仅有少量有标签样本数据甚至没有的学习问题，迁移学习被广泛应用于文本处理[7]、情感分类[8]、图像分类[9]、协同过滤[10]、医学[11]等领域，取得了很好的分类效果。

在网络异常检测基本领域，以前的方法只能检测已知的网络异常攻击，而迁移学习算法常常考虑了异常攻击和正常行为之间的潜在差异性。在对未知攻击类型进行决策时，可将这种潜在差异性迁移到当前的任务中。因此本文将迁移学习理论引入网络异常检测中，通过训练正常的数据和某种攻击来检测另外一种攻击。

由于网络异常的隐藏和伪装方式的不断进步，现阶段的网

收稿日期：2014-12-16；修回日期：2015-01-26 基金项目：国家自然科学基金资助项目(61103223)；江苏省自然科学基金重点资助项目(BK2011003)

作者简介：赵新杰(1990-)，男，江苏淮安人，硕士研究生，主要研究方向为网络安全、网络媒体智能技术(quintinzxj@gmail.com)；刘渊(1967-)，男，教授，硕导，主要研究方向为安全网络、数字媒体技术；孙剑(1989-)，男，硕士研究生，主要研究方向为网络安全、网络媒体技术及安全.

络异常可能只在某个特征方面和正常数据存在一定差异，所以本文将网络数据按照特征分成了连接特征、内容特征、流量特征三个部分，并且对每一部分都进行检测。为了使检测的结果更加精准和有效，必须选择一种数据融合的方法将三部分的结果融合在一起。在实际应用方面，网络异常检测中经常会出现离正常和异常两种行为都比较接近的情况，这样就造成了一种不确定性。通常的检测方法只是人工设置阈值，直接将这种不确定行为判定为其中的一种，而 D-S 理论可以为不确定信息的表达和合成提供强有力的理论依据。所以本文选择 D-S 理论对分类融合模型进行综合判决。

通过实验表明，本文提出的网络异常行为分析方法由于引入迁移学习方法，使得训练数据和测试数据的分布可以具有一定的差异性，即可以通过训练已知的网络异常行为来检测未知的网络异常行为，再运用 D-S 理论进行综合判决，能够提升检测的准确性和有效性。

1 迁移学习

1.1 迁移学习技术

迁移学习是运用已知的知识对不同但相关领域问题进行求解的一种新的机器学习方法。迁移学习不需要源域和目标域具有相同的分布，这样就解决了传统机器学习需要源域和目标域具有相同分布的局限性。

迁移学习大致分为三类：同构空间下基于实例的迁移学习、同构空间下基于特征的迁移学习和异构空间下的迁移学习。本文参考了文献[12]中提出的一种基于最大均值差(MMD)的直推式迁移学习方法(LMPROJ)。这种迁移学习算法的核心思想是寻找出某种特征变换的方法，使得基于某种样本分度距离度量的源域和目标域分布距离最小，同时使得训练样本的类与类之间的间隔最大化。

1.2 支持向量机(SVM)

由于 LMPROJ 算法是基于传统支持向量机的一种分类算法。SVM 算法[13]的基本思想是对于给定的两类样本的数据集 (x_i,y_i) $(i=1,2,\cdots,n, x\in \mathrm{R}^d, y\in\{+1,-1\})$，通过训练学习得到一个超平面 f 将其分为两类，使得分类间隔最大。在真正的分类实现过程中，引入了松弛变量到 SVM 的目标函数，最后形成了如下的数学表示形式：

$$f=\arg\min_{f\in H_K} C\sum_{i=1}^{n}V(x_i,y_i,f)+\frac{1}{2}\|f\|_K^2 \tag{1}$$

其中：K 为特征映射核函数 $K(x,x'):X\times X\to \mathrm{R}$；$H_K$ 是核空间之内的函数集；$\|f\|^2$ 是 f 的 L_2 范数形式；V 是用来预测训练样本类别的正则风险评估系数；C 是正规化系数，用来调节正则风险评估系数的结果。当 f 是一个线性函数时，用 w 表示式(1)如下：

$$\begin{aligned}&\min \frac{1}{2}\|w\|^2+C\sum_{j=1}^{n}\varepsilon_j\\&\text{s.t. } \varepsilon_j\geqslant 0, y_j(w^{\mathrm{T}}\varphi(x_j)+b)\geqslant 1-\varepsilon_j \quad j=1,\cdots,n\end{aligned} \tag{2}$$

1.3 大间隔直推式迁移学习方法

设源域 $D_s=\{\{(x_1,y_1),\cdots,(x_n,y_n)\}\}$，$x_j\in X_s$ 为源域的实际数据，$y_i\in Y_s$ 为相应的类别标签；目标域 $D_t=\{z_1,\cdots,z_m\}$，$z_k\in X_t$ 为目标域的实际数据，那么源域和目标域的最大均值差可以表示为

$$\mathrm{MMD}^2=\left\|\frac{1}{n}\sum_{j=1}^{n}\varphi(x_j)-\frac{1}{m}\sum_{k=1}^{m}\varphi(z_k)\right\|^2=\frac{1}{n^2}\sum_{j,k=1}^{n}K(x_j,x_k)+\frac{1}{m^2}\sum_{j,k=1}^{m}K(z_j,z_k)-\frac{2}{nm}\sum_{j,k=1}^{n,m}K(x_j,z_k) \tag{3}$$

LMPROJ 的目标函数构建基于 SVM 的思想，在结构风险最小的情况下，使得基于某种样本分度距离度量的源域和目标域分布距离最小，同时使得训练样本的类与类之间的间隔最大化，从而得到目标领域的预测函数：

$$F=\arg\min C\sum_{j=1}^{n}V(x_j,y_j,w)+\frac{1}{2}\|f\|_K^2+\lambda d_{f,k}(P_m,P_n)^2 \tag{4}$$

其中：P_m 为源域分布，P_n 为目标域分布，λ 为平衡参数，$d_{f,k}(P_m,P_n)^2$ 为两个样本领域的分布距离度量的平方。LMPROJ 只考虑线性决策 $f(x)=w^{\mathrm{T}}\varphi(x)$，$w$ 为投影向量，并定义投影的最大均值差距离度量来估计源域和目标域的距离。得到

$$d_{f,k}(P_m,P_n)^2=\left\|\frac{1}{n}\sum_{s}^{n}f(x_s)-\frac{1}{m}\sum_{t}^{m}f(z_t)\right\|^2=\frac{1}{n^2}\left(\sum_{s=1}^{n}w^{\mathrm{T}}\varphi(x_s)\right)^2+\frac{1}{m^2}\left(\sum_{t}^{m}w^{\mathrm{T}}\varphi(z_t)\right)^2-\frac{2}{nm}\left(\sum_{s=1}^{n}w^{\mathrm{T}}\varphi(x_s)\right)\left(\sum_{t}^{m}w^{\mathrm{T}}\varphi(z_t)\right) \tag{5}$$

根据上面给出的决策函数式(4)和距离函数式(5)，LMPROJ 的优化目标最终定义为

$$\begin{aligned}&\min \frac{1}{2}\|w\|^2+C\sum_{j=1}^{n}\varepsilon_j+\lambda d_{f,k}(P_m,P_n)^2\\&\text{s.t. } \varepsilon_j\geqslant 0, y_j(w^{\mathrm{T}}\varphi(x_j)+b)\geqslant 1-\varepsilon_j \quad \forall j=1,\cdots,n\end{aligned} \tag{6}$$

2 Dempster-Shafer 证据理论

2.1 D-S 理论基础

D-S 理论是建立在非空有限域 Θ 上的理论，Θ 成为辨识框架，即对某一问题所能认识到的所有可能的结果的集合。

定义 1 设 Θ 为辨识框架，如果集函数 $m:2^{\Theta}\to[0,1]$（2^{Θ} 为 Θ 的幂集）满足：a) $m(\varnothing)=0$；$\sum_{A\in\Theta}m(A)=1$。称 m 为辨识框架 Θ 上的信度分配函数(BPA)。

在信度函数的基础上又派生出了两个函数：信度函数(Bel)和似然函数(Pl)：

$$\begin{cases}\mathrm{Bel}(A)=\sum_{B\subset A}m(B)\\ \mathrm{Bel}(\varnothing)=0\\ \mathrm{Bel}(\Theta)=1\\ \mathrm{Pl}(A)=1-\mathrm{Bel}(\bar{A})=\sum_{B\cap A\neq\varnothing}m(B)\end{cases} \tag{7}$$

2.2 Dempster 合成法则

Dempster 的一般化规则为

$$\begin{cases}m_{1,\cdots,n}(A)=K^{-1}\sum_{\cap A_i=A}\prod_{1}^{n}m_i(A_i)\\ K=\sum_{\cap A_i\neq A}\prod_{1}^{n}m_i(A_i)\end{cases} \tag{8}$$

因为在文献[14]中已经证明，当 $\Theta=\{N,A\}$，$N\cap A=\varnothing$，即辨识框架为两个互斥元素时，Dempster 规则满足结合律：

$$m_{1,2,\cdots,n}(A)=m_{1,2,\cdots,n-1}(A)\oplus m_n(A) \tag{9}$$

所以本文只需要讨论两个规则的情况，设 m_1 和 m_2 分别是两个证据的信度分配函数，则

$$\begin{cases} m_1(A) \oplus m_2(A) = K^{-1} \sum_{A_1 \cap A_2 = A} m_1(A_1)\, m_2(A_2) \\ K = \sum_{A_1 \cap A_2 \neq A} m_1(A_1)\, m_2(A_2) \end{cases} \tag{10}$$

2.3　BPA 函数的确定

取辨识框架 Θ 为$\{N, A\}$，N 表示正常，A 表示攻击，考虑到 D-S 对冲突数据的局限性，定义 $m: P\{N, A\} \rightarrow [0, 1]$，$m(A) + m(N) + m(\Theta) = 1$。其中 $m(A)$ 表示属于攻击的可信度，$m(N)$ 表示属于正常数据的可信度，$m(\Theta)$ 表示不能确定属于正常还是攻击的可信度。

基于模式识别中相同类别之间的距离要小于不同类别之间的距离这一原则[15]，综合模糊数学中模糊隶属度的思想，BPA 函数构造如下（设 N 类的标签为 +，A 类的标签为 −）：

$$\begin{cases} m(N) = 1/2 + 1/2 \times (1 + \exp(-y))^{-1} \\ m(A) = 1/2 \times (1 - m(N)) \\ m(\Theta) = 1 - m(A) - m(N) \end{cases} \quad y > 0 \tag{11}$$

$$\begin{cases} m(A) = 1/2 + 1/2 \times (1 + \exp(y))^{-1} \\ m(N) = 1/2 \times (1 - m(A)) \\ m(\Theta) = 1 - m(A) - m(N) \end{cases} \quad y < 0$$

其中：y 是 LMPROJ 的输出。该 BPA 函数保证了当 LMPROJ 的输出 y 为正值时，属于 N 的概率大于属于 A 的概率，且 y 值越大属于 N 的概率越接近于 1；同理对于 y 为负值时属于 A 的概率大于属于 N 的概率且渐进为 1，符合相同类别之间的距离要小于不同类别之间的距离这一原则。

3　基于迁移算法和 D-S 理论的网络异常检测

基于迁移算法和 D-S 理论的网络异常检测模型，通过迁移学习的方法，使得训练数据和测试数据可以存在分布上的差异，弥补了以往网络异常检测中对未知行为无法有效检测的漏洞。并且由于 D-S 理论在不确定行为分析上的理论支持，不确定行为的判断更加有理论依据，最终检测的准确性和有效性都得到了提高。

3.1　基于迁移算法和 D-S 理论的网络异常检测模型

将数据集根据特征的不同分为三个部分，分别是连接特征、内容特征、流量特征，对每一个部分进行迁移学习，再将它们得到的分类结果标签运用 D-S 理论进行结合，得出最后的测试标签。基于迁移算法和 D-S 理论的网络异常检测模型如图 1 所示。

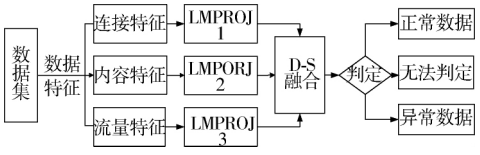

图 1　LMPROJ + D-S 的网络异常检测模型

3.2　基于迁移算法和 D-S 理论的网络异常检测的实现

输入：带有标签的 KDD99 数据集（正常数据标签为 1，攻击数据标签为 −1）。

输出：测试数据集的类别标签及实验的准确率和误报率。

a) 对 KDD99 的数据集分为五大类：正常数据（normal）、拒绝服务攻击（DoS）、远程主机未授权访问（R2L）、未授权的本地超级用户特权访问（U2R）、端口监视或扫描（Probing）。

b) 将每一类分成连接特征、内容特征和流量特征三类。

c) 每次使用正常数据和一种攻击构成训练数据集，利用 LMPROJ 对其训练，得到一个核分类器。

d) 将正常数据和另一种攻击组成测试数据集，用训练所得到的分类器对其进行测试，获得测试标签并生成各自的 BPA 函数值。

e) 将三种特征分别的 BPA 函数值运用 D-S 理论进行融合，得到最终的 $m(N)$、$m(A)$ 和 $m(\Theta)$ 的值。

f) 判断，当 $m(N)$ 为三者中最大时，将该数据认为是正常数据，设标签为 1；当 $m(A)$ 为三者中最大时，将该数据认为是攻击数据，设标签为 −1；当 $m(\Theta)$ 为三者中最大时，为了整个实验结果的准确性，设其标签为 0，这样判断与初始标签必然不同（即当做是错误的）。

g) 将测试标签和初始标签进行对比，得出实验的准确率。

4　实验

4.1　实验环境

本文采用的是 KDD99 数据集，该数据集来源于美国国防部高级规划署，并使用 Intel Core2 2.93GHz，2 GB 内存，Windows 7 操作系统的 PC 机在 MATLAB 2010 上进行仿真实验。

4.2　数据集预处理

由于 KDD99 是一个 41 维的数据集，其中既有连续型的数据，也有离散型和符号型的数据。对于符号型的数据，MATLAB 不能直接处理，所以本文通过一般映射将其映射成为离散型的数据，如第二维的 protocol_type（协议类型），分为 TCP、UDP、ICMP 三种，本文令 TCP = 1，UDP = 2，ICMP = 3。最后再使用 MATLAB 自带的 mapminmax 函数将数据集归一化。

4.3　实验分组

首先将 KDD99 数据集分成五组：A 组为正常数据、B 组为 R2L 攻击数据、C 组为 PROBING 攻击数据、D 组为 DoS 攻击数据、E 组为 U2R 攻击数据。每组数据再根据特征分成连接特征（1 ~ 9 维）、内容特征（10 ~ 22 维）和流量特征（23 ~ 41 维）三类。其中基本连接特征包含了一些连接的基本属性，如连续时间、协议类型、传送的字节数等；内容特征里抽取了部分可能反映入侵行为的内容特征，如登录失败的次数、访问系统敏感文件和目录的次数等；流量特征则包括了当前连接记录与之前一段时间内的连接记录之间存在的某些联系和过去两秒范围统计与当前连接之间的关系等。

为了验证本文方法的有效性，本文构造了两大类数据集来测试不同分类算法的性能：a) 训练样本和测试样本具有相同分布的；b) 训练样本和测试样本具有不同分布的。一共做了七组实验，其中 1、2 组是训练样本和测试样本具有相同分布的，3 ~ 7 组是训练样本和测试样本具有不同分布的。由于 KDD99 中 U2R 这类的攻击比较少，所以涉及到这组的实验，本文只选择了 200 个训练样本和 200 个测试样本，其他组的实验都选取了 500 个训练样本和 500 个测试样本，并且保证每组实验的训练样本和测试样本均不含有相同的部分，相互独立。实验分组情况如表 1 所示。

表 1 实验分组情况

分布特征	组别	训练数据集	测试数据集
训练测试	1	A、B 各 500	A、B 各 500
数据具有相同分布	2	A、C 各 500	A、C 各 500
训练测试	3	A、B 各 500	A、C 各 500
	4	A、B 各 500	A、D 各 500
数据具有不同分布	5	A、C 各 500	A、D 各 500
	6	A、C 各 500	A、E 各 500
	7	A、D 各 200	A、E 各 200

4.4 实验结果

表 2 中对每组实验中的每一类特征都运用了 LMPROJ 算法进行检测，检测的效果有好有坏，但是通过 D-S 理论对这三特征所得到的结果进行融合之后，得到的检测结果基本接近于三类特征检测中的最理想效果，具有较高的检测准确率。

表 2 各特征及融合后的检测正确率

实验组别	连接特征/%	内容特征/%	流量特征/%	D-S 融合后的结果/%
1	87.6	98.4	95.5	98.1
2	94.6	97.7	97.7	98.2
3	89.0	98.4	95.5	98.1
4	89.2	96.8	76.3	98.3
5	93.6	99.3	75.9	99.5
6	79.3	87.3	82.0	84.0
7	78.3	87.3	84.0	85.0

表 3 对每组实验的误报率进行了分析，可以很清楚地表明 D-S 理论对三个特征结果进行融合后的误报率都保持在一个较低的水准，而且可以看出基于迁移学习和 D-S 理论的网络异常检测整体有着较低的误报率。

表 3 各特征及融合后的检测误报率

实验组别	连接特征/%	内容特征/%	流量特征/%	D-S 融合后的结果/%
1	10.0	0.5	1.0	0.5`
2	1.0	1.5	0.4	0.9
3	9.4	1.4	2.8	0.9
4	11.6	1.9	1.6	1.1
5	5.1	0.6	2.8	0.5
6	3.5	1.0	1.3	0.8
7	4.3	0.2	5.0	1.7

表 4 表明运用 LMPROJ 算法和 D-S 理论融合的检测正确率比只使用 LMPROJ 算法的效果有了提升，可见 D-S 理论能够有效提升检测的准确率。另外图 2 也表明了本文的算法比 SVR 和 D-S 融合的算法效果提升了很多，甚至只使用迁移学习的算法就已经比 SVR 和 D-S 相融合的结果有所提升，验证了迁移学习对检测效果有明显的提高。

表 4 LMPROJ、SVR + D-S 和本文算法正确率的对比

实验组别	LMPROJ/%	SVR + D-S/%	LMPROJ + D-S/%
1	97.5	94.1	98.1
2	98.0	93.8	98.2
3	96.4	94.2	98.1
4	96.2	86.3	98.3
5	97.3	86.9	99.5
6	81.2	78.3	84.0
7	82.1	76.2	85.1

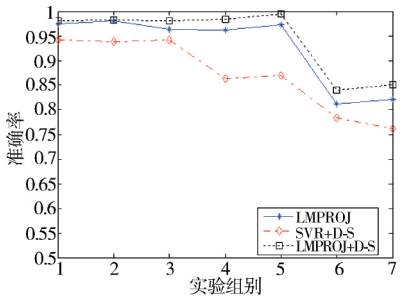

图 2 LMPROJ、SVR + D-S 和本文的 ROC 曲线图

5 结束语

迁移学习作为一种比较新的理论，已经运用在很多的领域。本文探讨了其在网络异常检测方面的应用，研究了一种基于迁移学习和 D-S 证据理论的网络异常检测方法，并使用 KDD99 数据集进行了实验，得出的实验结果表明，基于迁移学习和 D-S 证据理论的方法比一般的分类算法和 D-S 证据理论相结合或者是单独使用迁移学习算法的结果都要理想，具有一定的优越性。随着对迁移学习的不断理解和深入发展，相信迁移学习在网络异常检测方面肯定还拥有更广阔的发展空间。

参考文献:

[1] Staniford S, Hoagland J A, McAlerney J M. Practical automated detection of stealthy portscans [J]. Journal of Computer Security, 2002, 10(1): 105-136.

[2] 郁继锋. 基于数据挖掘的 Web 应用入侵异常检测研究 [D]. 武汉: 华中科技大学, 2011.

[3] 胡明霞. 基于 BP 神经网络的入侵检测算法 [J]. 计算机工程, 2012, 38(6): 148-150.

[4] 张剑, 龚俭. 一种基于模糊综合评判的入侵异常检测方法 [J]. 计算机研究与发展, 2003, 40(6): 776-783.

[5] 黄学宇, 魏娜, 陶建锋. 基于人工免疫聚类的异常检测算法 [J]. 计算机工程, 2010, 36(1): 166-169.

[6] 庄福振, 何清, 史忠植. 迁移学习研究进展 [J]. 软件学报, 2015, 26(1): 26-39.

[7] Dai Wenyuan, Xue Guirong, Yang Qiang, *et al*. Co-clustering based classification for out-of-domain documents [C] //Proc of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2007: 210-219.

[8] Li Tao, Zhang Yi, Sindhwani V. A non-negative matrix tri-factorization approach to sentiment classification with lexical prior knowledge [C] //Proc of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing. 2009: 244-252.

[9] Raina R, Battle A, Lee H, *et al*. Self-taught lear ning: transfer learning from unlabeled data [C] //Proc of the 24th International Conference on Machine Learning. New York: ACM Press, 2007: 759-766.

[10] Cao Bin, Liu N N, Yang Qiang. Transfer learning for collective link prediction in multiple heterogenous domains [C] //Proc of the 27th International Conference on Machine Learning. 2010: 159-166.

[11] 杨昌健, 邓赵红, 蒋亦樟, 等. 基于迁移学习的癫痫 EEG 信号自适应识别 [J]. 计算机科学与探索, 2014, 8(3): 329-337.

[12] Quanz B, Huan Jun. Large margin transductive transfer learning [C] // Proc of the 18th ACM Conference on Information and Knowledge Management. New York: ACM Press, 2009: 1327-1336.

[13] 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000, 26(1): 36-46.

[14] 诸葛建伟, 王大为, 陈昱, 等. 基于 D-S 证据理论的网络异常检测方法 [J]. 软件学报, 2006, 17(3): 463-471.

[15] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究 [J]. 软件学报, 2008, 19(1): 48-61.