基于粒子群优化多核支持向量数据描述的广播式自动相关监视异常数据检测模型

王布宏 罗 鹏* 李腾耀 田继伟 尚福特

(空军工程大学 信息与导航学院 西安 710077)

摘 要：广播式自动相关监视(ADS-B)作为新一代空中交通管理(ATM)通信协议，是未来空管监视系统的关键技术。目前，由于ADS-B采用明文格式广播发送数据，其安全性问题受到挑战。针对ADS-B易受到的欺骗干扰，该文将ADS-B位置数据和同步的二次雷达(SSR)数据作差，将两者的差值作为样本数据。利用多核支持向量数据描述(MKSVDD)训练样本，得到了超球体分类器，此超球体分类器能检测出ADS-B测试样本中的异常数据。并且，通过粒子群算法(PSO)优化了GaussLapl和GaussTanh两种MKSVDD的惩罚因子、多核核函数系数以及核参数，提高了异常数据检测性能。实验结果表明，对于随机位置偏移、固定位置偏移、拒绝服务(DOS)攻击和重放攻击，粒子群优化多核支持向量数据描述(PSO-MKSVDD)模型能检测出这4种攻击类型的异常数据。且相较于其他机器学习和深度学习方法，该模型的适应性更好，异常检测的召回率和检测率更优。证明该模型可用于ADS-B异常数据的检测。

关键词：广播式自动相关监视；空中交通管理；异常检测；多核支持向量数据描述；粒子群优化

中图分类号：TN967.1; TP391 **文献标识码**：A **文章编号**：1009-5896(2020)11-2727-08

DOI: 10.11999/JEIT190767

ADS-B Anomalous Data Detection Model Based on PSO-MKSVDD

WANG Buhong LUO Peng LI Tengyao TIAN Jiwei SHANG Fute

(*School of Information and Navigation, Air Force Engineering University, Xi'an* 710077, *China*)

Abstract: As a new generation of Air Traffic Management(ATM) communication protocol, Automatic Dependent Surveillance-Broadcast(ADS-B) is the key technology of ATM monitoring system in the future. At present, the security of ADS-B is challenged because it broadcasts data in plaintext format. Because ADS-B is susceptible to spoofing, the difference between ADS-B position data and synchronous Secondary Surveillance Radar(SSR) data is taken as sample data. Using Multi-Kernel Support Vector Data Description(MKSVDD) to train samples, a hypersphere classifier is obtained, which can detect anomalous data in ADS-B test samples. In addition, Particle Swarm Optimization (PSO) is used to optimize GaussLapl and GaussTanh MKSVDD penalty factors, coefficients of multi-kernel functions and kernel parameters.The performance of anomaly detection is improved. Experimental results show that PSO-MKSVDD can detect anomalous data of random position deviation, fixed position deviation, Denial Of Service(DOS) attack and replay attack. In addition, compared with other machine learning and deep learning methods, this model has better adaptability and better recall rate and detection rate of anomaly detection.It is proved that this model can be used to detect ADS-B anomalous data.

Key words: Automatic Dependent Surveillance-Broadcast (ADS-B); Air Traffic Management (ATM); Anomaly detection; Multi-Kernel Support Vector Data Description (MKSVDD); Particle Swarm Optimization (PSO)

收稿日期：2019-10-08；改回日期：2020-04-04；网络出版：2020-04-29

*通信作者： 罗鹏 1939552724@qq.com

基金项目：国家自然科学基金(61902426)

Foundation Item: The National Natual Science Foundation of China (61902426)

1 引言

空中交通管理(Air Traffic Management, ATM)是管理空域、飞行器和航线的重要系统。它的任务是保护和促进空中交通安全，维护空中交通秩序，保障空中交通畅通。目前，随着空中交通流量的持续增大，导致了空中交通的拥塞，这给空中交通管制(Air Traffic Control, ATC)带来了极大的困难。一种有效的解决方案是提升空管监视系统的定位精度，传统的空管监视技术主要包括一次雷达(Primary Surveillance Radar, PSR)、二次雷达(Secondary Surveillance Radar, SSR)以及多点定位(Multi-LATeration, MLAT)，但雷达的监测精度相对较低、多点定位对地面基站的安装位置有一定要求，因而广播式自动相关监视(Automatic Dependent Surveillance-Broadcast, ADS-B)技术正在发挥主要作用[1]。ADS-B通过全球卫星导航系统(Global Navigation Satellite System, GNSS)获取其位置信息、通过机载导航设备获取其速度和航向等信息，而后通过ADS-B发射器广播这些信息，其他装有ADS-B接收设备的飞机和地面基站能接收到这些信息。作为新一代空管监视系统的关键技术，ADS-B定位精度达到33 m，比雷达200 m精度有大幅提升[2]。由于ADS-B设备经济成本低和监视范围更加全面，全球大部分航空公司将于2020年前完成部署ADS-B设备。然而，由于ADS-B以广播方式发送明文数据，协议没有提供消息认证和加密机制，因而ADS-B安全性难以保证。

文献[3]讨论了ADS-B易遭受的窃听、干扰、消息注入、消息删除和消息篡改等攻击，并分析了攻击的难易等级。文献[4]从网络分层、攻击方法和影响程度等角度分析了ADS-B遭受的攻击。针对ADS-B的安全漏洞，目前的解决方案主要包括以下3类。第1类方案将密码学的方案应用于ADS-B[5,6]，但是这类方案存在一些困难和问题：(1)ADS-B是一项国际协议，加解密方案必须和现有的标准化协议相适应；(2)一旦密钥管理出现问题，如密钥泄露给ADS-B非法用户，则系统的安全性仍将遭受损害。第2类方案是将位置认证用于ADS-B数据的合法性判定[7]。文献[8]综合运用到达时间差(Time Difference Of Arrival, TDOA)、到达角(Angle Of Arrival, AOA)和到达频率差(Frequency Difference Of Arrival, FDOA)等量测定位飞行器位置，和ADS-B解析后的位置数据进行比较，实现ADS-B合法性判定，但是这对地面基站的安装位置有一定要求，并且需要多个基站协同工作，经济成本较高。文献[9]利用ADS-S消息对飞机的运动性能进行建模，利用建立的模型确定飞机不同飞行阶段的活动范围，进而判定ADS-B位置数据的合法性。第3类方案将机器学习和深度学习方法用于ADS-B数据的合法性判定。此类方案通常用于ADS-B数据的异常检测，异常检测是ADS-B研究的一个非常重要的问题。文献[10,11]分别采用深度学习的LSTM-encoder-decoder和seq2seq模型对ADS-B报文数据进行重构，通过重构误差检测ADS-B数据中的位置和速度等异常，但该方法没有考虑飞机的机动性状态带来的ADS-B数据变化，此方法的适应性较差。文献[12]将支持向量数据描述(Support Vector Data Description, SVDD)机器学习算法运用于ADS-B位置数据的异常检测，但是此方法没有考虑ADS-B目标类数据的多数据源和异构数据集问题，异常检测的检测率较低，异常检测性能较差。

本文针对ADS-B易遭受的随机位置偏移、固定位置偏移、拒绝服务(Denial Of Service, DOS)攻击和重放攻击，构建了粒子群优化多核支持向量数据描述(Particle Swarm Optimization Multi-Kernel Support Vector Data Description, PSO-MKSVDD)异常检测模型，用以检测以上4种攻击类型的异常数据。并且，此模型考虑了飞机的机动性状态带来的ADS-B数据变化，比文献[10,11]适应性更好；考虑ADS-B目标类数据的多数据源和异构数据集问题，比文献[12]ADS-B异常检测的检测率更优。本文首先介绍了SVDD、多核支持向量数据描述(Multi-Kernel Support Vector Data Description, MKSVDD)和粒子群算法(Particle Swarm Optimization, PSO)。接着，将ADS-B位置数据和SSR数据作差，把差值作为样本，样本经过训练后，得到了可用于异常检测的超球体分类器。并且，利用PSO算法优化了惩罚因子、多核核函数系数以及MKSVDD核参数，进一步提高了异常检测的召回率和检测率性能。仿真实验结果表明，PSO-MKSVDD能检测出这4种攻击类型的异常数据。

2 准备知识

2.1 支持向量数据描述

支持向量数据描述(SVDD)是Tax等人[13]提出的一种单分类无监督机器学习方法。其目的是寻找一个几乎包含全部训练样本的最小超球体。对于训练集$X=\{\boldsymbol{x}_i\in\boldsymbol{R}^d\,|i=1,2,\cdots,n\}$, SVDD表示为求解如式(1)所示问题

$$\begin{aligned}&\min\ \ F(R,\boldsymbol{a})=R^2+C\sum_{i=1}^{n}\xi_i,\\&\text{s.t.}\ \begin{cases}\|\boldsymbol{x}_i-\boldsymbol{a}\|^2\le R^2+\xi_i, i=1,2,\cdots,n\\ \xi_i\ge 0\end{cases}\end{aligned}\tag{1}$$

其中，$\boldsymbol{x}_i$是d维向量，R为半径，$\boldsymbol{a}$为球心。ξ_i为松

弛因子，$\xi_i \geq 0$。C是惩罚系数，用来控制超球体体积和非目标类的数量。引入拉格朗日乘子$\alpha_i(\alpha_i \geq 0)$和$\beta_i(\beta_i \geq 0)$，可求解出超球体半径$R$满足

$$R^2 = (\boldsymbol{x}_k, \boldsymbol{x}_k) - 2\sum_{i=1}^{n}\alpha_i(\boldsymbol{x}_i, \boldsymbol{x}_k) + \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{2}$$

则测试样本$\boldsymbol{x}_t$与球心$\boldsymbol{a}$的距离$f(\boldsymbol{x}_t)$满足

$$f^2(\boldsymbol{x}_t) = \| \boldsymbol{x}_t - \boldsymbol{a}\|^2 = (\boldsymbol{x}_t, \boldsymbol{x}_t) - 2\sum_{i=1}^{n}\alpha_i(\boldsymbol{x}_t, \boldsymbol{x}_i) + \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{3}$$

若$f(\boldsymbol{x}_t) < R$，则测试样本$\boldsymbol{x}_t$在超球体内部，是正常样本；若$f(\boldsymbol{x}_t) = R$，则测试样本$\boldsymbol{x}_t$是支持向量；若$f(\boldsymbol{x}_t) > R$，则测试样本$\boldsymbol{x}_t$是异常样本。

2.2 多核核函数

在求解超球体半径R和距离$f(\boldsymbol{x}_t)$时，为了使分类结果更具鲁棒性，需要用核函数代替内积

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{4}$$

但不同的核函数针对不同类型的数据集，具有不同的分类效果。MKSVDD能解决目标类数据存在多数据源和异构数据集的分类问题。多核核函数的构造包括多尺度核、无限核和合成核方法[14,15]。合成核中的加权求和核方法通过调节权重系数，能使多核核函数适应不同的样本输入，故这种方法兼具学习与泛化能力；且该方法的运算复杂度低，故该文采用两两加权求和核方法构造多核核函数

$$K_{\text{multi}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = aK_1(\boldsymbol{x}_i, \boldsymbol{x}_j) + bK_2(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{5}$$

其中，a和b是系数，满足$a > 0$, $b > 0$, $a + b = 1$; $K_1(\boldsymbol{x}_i, \boldsymbol{x}_j)$和$K_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$是单核函数。常见的单核函数有Gauss核函数$K_{\text{Gauss}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$, Sigmoid核函数$K_{\text{Tanh}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$以及Laplace核函数$K_{\text{Lapl}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$，具体形式为

$$\left.\begin{aligned} K_{\text{Gauss}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= \exp(-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2/{s_1}^2) \\ K_{\text{Tanh}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= \tanh(\beta{\boldsymbol{x}_i}^{\text{T}}\boldsymbol{x}_j + \theta) \\ K_{\text{Lapl}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= \exp(-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2/s_2) \end{aligned}\right\} \tag{6}$$

本文构造3种多核核函数，分别为GaussLapl多核核函数$K_{\text{GL}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$, GaussTanh多核核函数$K_{\text{GT}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$以及TanhLapl多核核函数$K_{\text{TL}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$，具体形式分别为

$$\left.\begin{aligned} K_{\text{GL}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= aK_{\text{Gauss}}(\boldsymbol{x}_i, \boldsymbol{x}_{mj}) + bK_{\text{Lapl}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ K_{\text{GT}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= aK_{\text{Gauss}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + bK_{\text{Tanh}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ K_{\text{TL}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= aK_{\text{Tanh}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + bK_{\text{Lapl}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \end{aligned}\right\} \tag{7}$$

2.3 PSO算法机理

粒子群优化算法(Particle Swarm Optimization, PSO)是一种基于群智能的演化计算技术[16]。目前PSO算法已经广泛应用于函数优化、神经网络训练等领域[17]。待优化参数被称作粒子，粒子具有两个属性：速度和位置。所有粒子均由适应度函数决定它的适应值。本文适应值由异常检测的错误率决定。每次迭代中，粒子根据适应值大小搜索自身的最优解，将其记为个体极值。整个种群搜索得到的最优解记为全局极值。粒子根据个体极值和全局极值来调整自身的速度和位置。将第i个粒子的速度表示为$\boldsymbol{v}_i$，位置表示为$\boldsymbol{x}_i$；粒子总数表示为N；第i个粒子的个体极值表示为$\mathbf{pbest}_i$；若在所有的个体极值中，第g个最优，则将其作为全局极值并表示$\mathbf{gbest}_g$; rand_1和rand_2表示介于(0,1)之间的均匀分布的随机数；速度和位置的更新公式为

$$\boldsymbol{v}_{i+1} = w\boldsymbol{v}_i + c_1\text{rand}_1(\mathbf{pbest}_i - \boldsymbol{x}_i) + c_2\text{rand}_2(\mathbf{gbest}_g - \boldsymbol{x}_i) \tag{8}$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{v}_i \tag{9}$$

其中，$1 \leq i \leq N$; w是惯性权重,用来衡量粒子搜索能力的大小；c_1是个体学习因子；c_2是群体学习因子；一般取$c_1 = c_2 = 2$。本文采用PSO算法优化惩罚因子、多核核函数系数以及MKSVDD核参数。

3 异常检测模型构建

3.1 特征数据选取

ADS-B协议中包含了飞机的经度、纬度、高度、识别号、速度和航向等信息，本文针对ADS-B位置数据出现的异常，构建了ADS-B异常检测模型。由于ADS-B位置数据随时间的变化而变化，不同时间的ADS-B数据不能保证分布在超球体内部的，因而不能直接对ADS-B位置数据直接用MKSVDD算法。本文将ADS-B位置数据和同步的SSR数据作差，将两者的差值$\boldsymbol{x}(k)$选取作为MKSVDD异常检测的特征样本数据。对于同一空中目标的监视，ADS-B的精度约33 m，SSR的监视精度约为200 m，所以，无论飞机处于何种机动状态，两者的差值能保证收敛在一个常数H范围内

$$\left.\begin{aligned} -\text{H} < \boldsymbol{x}(k) < \text{H} \\ \boldsymbol{x}(k) = \mathbf{ads}(k) - \mathbf{ssr}(k) \end{aligned}\right\} \tag{10}$$

3.2 PSO-MKSVDD异常检测模型

PSO-MKSVDD模型旨在将PSO优化算法和MKSVDD结合起来，用以检测ADS-B异常数据。图1是ADS-B异常检测模型的流程图。模型包括两个阶段：PSO参数优化阶段和MKSVDD异常检测阶段。

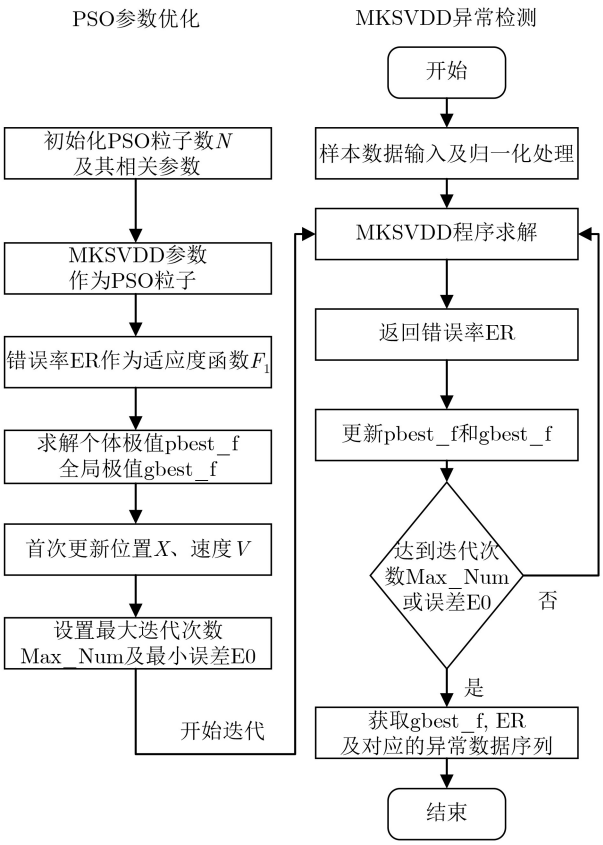

图 1 PSO-MKSVDD异常检测模型

3.2.1 PSO参数优化阶段

步骤 1 初始化PSO粒子数N、惯性权重w、个体学习因子c_1和群体学习因子c_2。

步骤 2 MKSVDD参数作为PSO粒子。MKSVDD参数包括惩罚因子C、多核核系数a, b以及MKSVDD核参数。将这些待优化参数视作粒子，利用PSO算法优化这些参数。

步骤 3 错误率ER作为PSO适应度函数F_1。错误率ER用来描述异常检测分类器的错分比例

$$\mathrm{ER}=(\mathrm{FP}+\mathrm{FN})/(P_1+N_1) \tag{11}$$

P_1是正样本数量，N_1是负样本数量；FP是被模型预测为正的负样本数量；FN是被模型预测为负的正样本数量。待优化粒子经过适应度函数计算，产生适应值。粒子的适应值是更新自身个体极值的依据。

步骤 4 求解个体极值**pbest_f**和全局极值**gbest_f**。粒子根据适应值大小迭代搜索自身的最优解，记为个体极值**pbest_f**；全局极值**gbest_f**为所有个体极值中最小值。

步骤 5 首次更新位置X、速度V。粒子根据式(8)和式(9)更新自己的位置X和速度V。

步骤 6 设置最大迭代次数Max_Num及最小误差E0。当迭代次数达到Max_Num时，粒子群停止迭代搜索。当错误率ER小于E0，粒子群同样会停止迭代搜索。

3.2.2 MKSVDD异常检测阶段

步骤 1 样本数据选取及归一化处理。将$\boldsymbol{x}(k)=\mathbf{ads}(k)-\mathbf{ssr}(k)$选取样本数据，$\boldsymbol{x}(k)$是一个3维列向量，包含纬度、经度和高度信息。样本数据的训练，需要先作归一化处理，归一化处理公式为

$$\boldsymbol{x}(k)=((\boldsymbol{x}(k)-\mu)/\sigma) \tag{12}$$

其中，μ是样本的均值，σ是样本的标准差。

步骤 2 MKSVDD程序求解。样本数据经过训练，得到超球体半径R和样本$\boldsymbol{x}_t$到球心的距离$f(\boldsymbol{x}_t)$。通过比较$f(\boldsymbol{x}_t)$和R的大小，能判断样本$\boldsymbol{x}_t$的合法性。

步骤 3 返回错误率ER。根据MKSVDD程序求解出的错误分类的样本数量占全部测试样本的比例，得到错误率ER。

步骤 4 更新**pbest_f**和**gbest_f**。每个粒子通过比较当前的适应值与历史个体极值的大小，来更新个体极值。若当前的适应值小于历史个体极值，则个体极值更新为当前的适应值；否则个体极值仍取值为历史个体极值。全局极值取值为所有个体极值中最小值。

步骤 5 达到最大迭代次数Max_Num或最小误差E0。如果达到设定的最大迭代次数，或者错误率小于误差E0时，则终止迭代。否则，转到步骤2，继续迭代搜索。

步骤 6 获取**gbest_f**, ER及对应的异常数据序列。全局极值**gbest_f**是用PSO算法求出的最优参数；错误率ER可以初步评估PSO-MKSVDD异常检测模型的整体性能；异常数据序列是PSO-MKSVDD模型求解出的非法数据。

3.3 异常检测模型评价指标

本文采用召回率Re和检测率De评估异常检测模型的性能，召回率Re是指被模型正确预测为正样本的数量占实际正样本数的比例，检测率De是指被模型正确预测出为负样本的数量占实际负样本数的比例。召回率Re和检测率De的值越高，PSO-MKSVDD异常检测模型的性能就越佳。表1是样本的分类结果。

召回率Re和检测率De分别定义为

表 1 样本分类结果表

实际情况	预测结果	
	正例	负例
正例	TP(真正例)	FN(假负例)
负例	FP(假正例)	TN(真负例)

$$\left.\begin{aligned}\mathrm{Re}&=\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})\\ \mathrm{De}&=\mathrm{TN}/(\mathrm{FP}+\mathrm{TN})\end{aligned}\right\}\tag{13}$$

4　仿真与实验结果

4.1　数据获取

实验从OPENSKY中获取200架次航班数据[18]，每架次ADS-B数据在100～1000之间，选取的ADS-B数据包含了飞机的起飞、爬升、巡航和下降阶段。同步的SSR数据从2次雷达监视系统数据库中获得。图2—图5，是一次包含300条ADS-B数据的航班，飞机处于爬升到巡航的转换阶段，此时飞机的机动状态变化较大。实验中，将前200条航迹数据用作训练样本，可以训练出超球体分类器；后100条航迹数据用作测试样本。攻击的报文数据由模拟生成，具体生成方法如下：

随机位置偏移(消息篡改)：如图2所示，共300条数据构成了ADS-B航迹。其中，用于测试的后100条ADS-B航迹数据中：前50条数据不做任何篡改；后50条数据，实验在经度和纬度上加上高斯白噪声，高斯白噪声的均值为0，标准差为800 m。

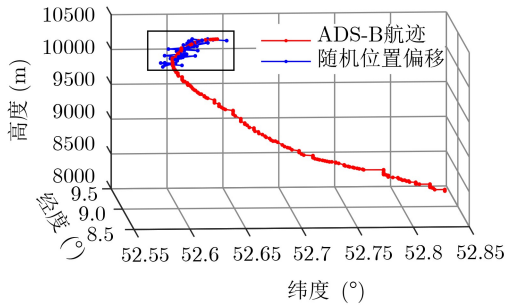

图 2 随机位置偏移攻击

固定位置偏移(消息篡改)：如图3所示，前50条测试航迹数据不做篡改；后50条数据，在经度和纬度上加上400 m的固定数值偏差。

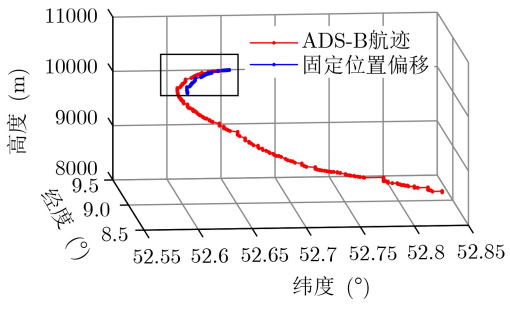

图 3 固定位置偏移攻击

航迹消失(DOS攻击)：如图4所示，前50条测试航迹数据不做篡改；后50条数据，由于受到DOS攻击，航迹数据不能被监视到，出现了航迹消失。

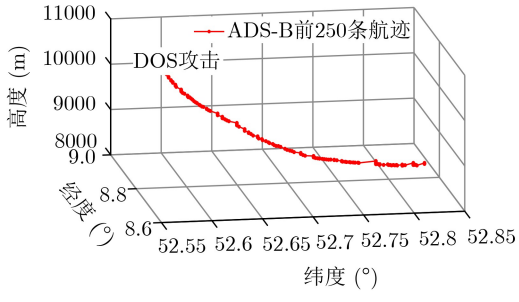

图 4 DOS攻击

航迹替换(重放攻击)：如图5所示，前50条测试航迹数据不做篡改；针对后50条数据，攻击者向监视系统发送的并非真实航迹，而是延迟发送的前50条测试航迹。

如式(11)所示，错误率表示分类器的错分比例，一般来说错误率越低，分类效果越好，异常检测性能越佳。实验中，将异常检测的错误率ER作为PSO算法的适应度函数。实验分别对单核SVDD，以及GaussLapl, GaussTanh和TanhLapl这3种MKSVDD的参数进行优化。如图6—图9所示，当异常检测的错误率最低的时候，单核SVDD的惩罚因子C为0.25; GaussLapl惩罚因子C为0.37；GaussTanh惩罚因子C为0.28; TanhLapl惩罚因子C为0.16。

用PSO算法优化其它参数。SVDD的核参数s_1=9。GaussLapl核参数s_1=5, s_2=20；系数a=0.8, b=0.2。GaussTanh的s_1=5, β=0.1, θ=10; a=0.4, b=0.6。TanhLapl的s_2=10, β=0.2, θ=1; a=0.7, b=0.3。

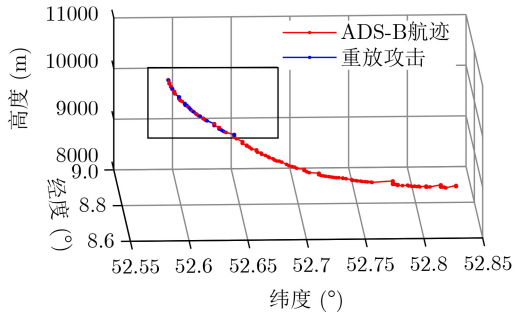

图 5 重放攻击

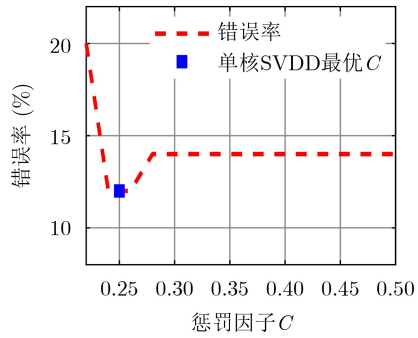

图 6 单核SVDD最优惩罚因子C

4.2 异常检测实验结果

针对随机位置偏移、固定位置偏移、DOS攻击和重放攻击，实验分别比较了单核SVDD, GaussLapl, GaussTanh和TanhLapl的异常检测的召回率与检测率。

图10是单核SVDD的随机位置偏移的检测图，样本经过训练后，得到的超球体半径R为0.52。经计算，SVDD随机位置偏移的召回率是94%，检测率是90%。如图11—图13所示，GaussLapl随机位置偏移的召回率是98%，检测率是94%。GaussTanh召回率是96%，检测率是92%。TanhLapl的召回率是96%，检测率是86%。由于TanhLapl的检测率相对较差，所以后续实验不再考虑TanLapl。

针对单核SVDD以及GaussLapl, GaussTanh两种MKSVDD，将200架次航班数据做了异常检测实验，取200次实验召回率和和检测率的均值作为实验的结果。表2综合对比了SVDD, GaussLapl, GaussTanh在4种不同攻击下，异常检测的召回率和检测率的值。对比分析表2中的数据，结果表

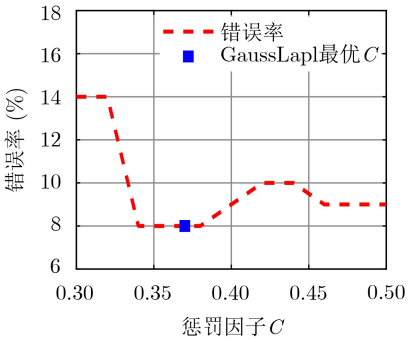

图 7 GaussLapl最优惩罚因子C

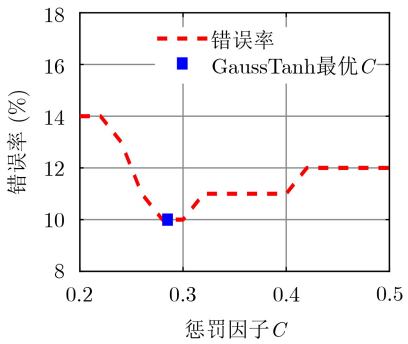

图 8 GaussTanh最优惩罚因子C

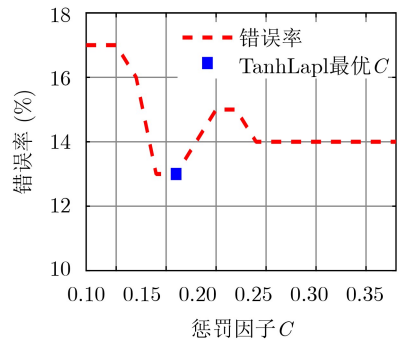

图 9 TanhLapl最优惩罚因子

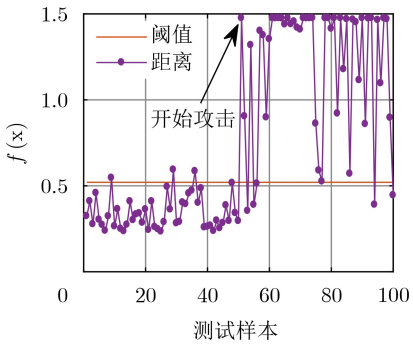

图 10 单核SVDD异常检测

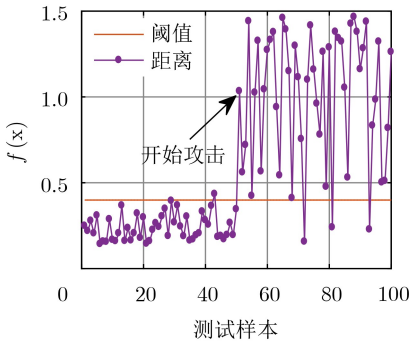

图 11 GaussLapl异常检测

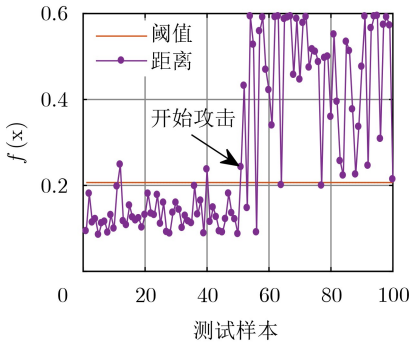

图 12 GaussTanh异常检测

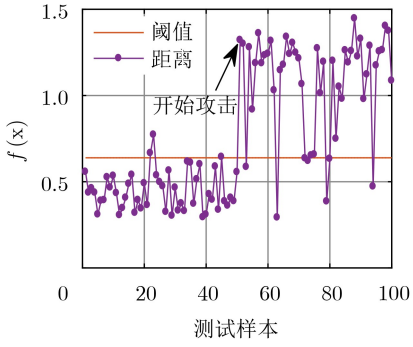

图 13 TanhLapl异常检测

表 2　异常检测对比表(%)

		SVDD	GaussLapl	GaussTanh
随机位置偏移	召回率	94.0	95.2	94.8
	检测率	89.2	93.6	92.0
固定位置偏移	召回率	94.8	95.6	96.0
	检测率	94.4	96.4	97.2
DOS攻击	召回率	94.8	96.0	95.2
	检测率	100.0	100.0	100.0
重放攻击	召回率	94.8	96.0	95.6
	检测率	98.4	99.2	98.9

明：GaussLapl和GaussTanh方法的召回率与检测率均高于单核SVDD，这证明了本文PSO-MKSVDD异常数据检测模型的正确性和优势。

另外，本文选取了其他机器学习和深度学习的方法，与PSO-MKSVDD模型作比较。这些方法包括长短期记忆神经网络(Long Short Term Memory, LSTM), SVDD[12], LSTM-encoder-decoder[10]和seq2seq[11]。对比实验中，ADS-B攻击数据的生成方法和4.1节的4种攻击方式保持一致(即在用相同的ADS-B攻击数据条件下，比较了各种异常检测方法的性能)。表3列出了各种方法结果对比表。对比分析表3中的数据，结果表明：

(1) GaussLapl和GaussTanh方法的召回率和检测率，均比LSTM-encoder-decoder和seq2seq的要高。这是因为本文考虑了飞机机动性状态带来的ADS-B数据变化，因而PSO-MKSVDD异常检测性能更优。

表 3　各种异常检测方法结果对比(%)

		LSTM	SVDD	LSTM-encoder-decoder	seq2seq	GaussLapl	GaussTanh
随机位置偏移	召回率	85.6	94.0	90.3	91.7	95.2	94.8
	检测率	87.0	89.2	89.8	90.6	93.6	92.0
固定位置偏移	召回率	84.2	94.8	93.8	91.0	95.6	96.0
	检测率	72.1	94.4	79.4	82.4	96.4	97.2
DOS攻击	召回率	87.5	94.8	93.7	94.4	96.0	95.2
	检测率	92.6	100.0	95.2	95.6	100.0	100.0
重放攻击	召回率	85.7	94.8	92.0	91.6	96.0	95.6
	检测率	88.2	98.4	93.6	94.4	99.2	98.9

(2) 此模型比SVDD用于ADS-B异常数据检测的检测率和召回率更高。这是因为PSO-MKSVDD模型考虑了ADS-B数据的异构性和复杂性，并且用PSO算法优化了实验参数。

5　结论

本文将ADS-B位置数据和同步的SSR数据作差，把差值作为MKSVDD的样本，样本经过训练后得到了超球体分类器，此分类器可用于检测ADS-B随机位置偏移、固定位置偏移、DOS攻击和重放攻击。且用PSO算法优化了惩罚因子、多核核函数系数以及核参数，进一步提高了异常检测的召回率和检测率。

PSO-MKSVDD模型考虑了飞机机动性状态带来的ADS-B数据变化，比深度学习的LSTM-encoder-decoder和seq2seq模型的适应性更好，异常检测性能更优；PSO-MKSVDD模型考虑了ADS-B数据的异构性和复杂性，比SVDD的异常数据召回率和检测率更高，异常检测性能更佳。

参 考 文 献

[1] STROHMEIER M, SCHAFER M, LENDERS V, *et al*. Realities and challenges of nextgen air traffic management: The case of ADS-B[J]. *IEEE Communications Magazine*, 2014, 52(5): 111–118. doi: 10.1109/MCOM.2014.6815901.

[2] ZHANG Jun, LIU Wei, and ZHU Yanbo. Study of ADS-B data evaluation[J]. *Chinese Journal of Aeronautics*, 2011, 24(4): 461–466. doi: 10.1016/s1000-9361(11)60053-8.

[3] MCCALLIE D, BUTTS J, and MILLS R. Security analysis of the ADS-B implementation in the next generation air transportation system[J]. *International Journal of Critical Infrastructure Protection*, 2011, 4(2): 78–87. doi: 10.1016/j.ijcip.2011.06.001.

[4] STROHMEIER M, LENDERS V, and MARTINOVIC I. On the security of the automatic dependent surveillance-broadcast protocol[J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(2): 1066–1087. doi: 10.1109/comst.2014.2365951.

[5] WESSON K D, HUMPHREYS T E, and EVANS B L. Can cryptography secure next generation air traffic surveillance?[J/OL]. IEEE Security & Privacy. http://radionavlab.ae.utexas.edu/images/stories/files/papers/adsb_for_submission.pdf, 2014.

[6] BAEK J, BYON Y J, HABLEEL E, *et al*. Making air traffic surveillance more reliable: A new authentication framework

for automatic dependent surveillance-broadcast (ADS-B) based on online/offline identity-based signature[J]. *Security and Communication Networks*, 2015, 8(5): 740–750. doi: 10.1002/sec.1021.

[7] MONTEIRO M. Detecting malicious ADS-B broadcasts using wide area multilateration[C]. The IEEE/AIAA 34th Digital Avionics Systems Conference, Prague, Czekh, 2015. doi: 10.1109/DASC.2015.7311579.

[8] NIJSURE Y A, KADDOUM G, GAGNON G, *et al*. Adaptive air-to-ground secure communication system based on ADS-B and wide-area multilateration[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(5): 3150–3165. doi: 10.1109/TVT.2015.2438171.

[9] SUN J, ELLERBROEK J, and HOEKSTRA J M. Modeling aircraft performance parameters with open ADS-B data[C]. The 12th USA/Europe Air Traffic Management Research and Development Seminar, Seattle, USA, 2017.

[10] HABLER E and SHABTAI A. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B Messages[J]. *Computers & Security*, 2018, 78: 155–173. doi: 10.1016/j.cose.2018.07.004.

[11] 丁建立, 邹云开, 王静, 等. 基于深度学习的ADS-B异常数据检测模型[J]. 航空学报, 2019, 40(12): 323220. doi: 10.7527/S1000-6893.2019.23220.
DING Jianli, ZOU Yunkai, WANG Jing, *et al*. ADS-B anomaly data detection model based on deep learning[J]. *Acta Aeronautica et Astronautica Sinica*, 2019, 40(12): 323220. doi: 10.7527/S1000-6893.2019.23220.

[12] 王振昊, 王布宏. 基于SVDD的ADS-B异常数据检测[J]. 河北大学学报: 自然科学版, 2019, 39(3): 323–329. doi: 10.3969/j.issn.1000-1565.2019.03.015.
WANG Zhenhao and WANG Buhong. ADS-B anomaly data detection based on SVDD[J]. *Journal of HeBei University*: *Natural Science Edition*, 2019, 39(3): 323–329. doi: 10.3969/j.issn.1000-1565.2019.03.015.

[13] TAX D M J and DUIN R P W. Support vector data description[J]. *Machine Language*, 2004, 54(1): 45–66. doi: 10.1023/b:Mach.0000008084.60811.49.

[14] GÖNEN M, and ALPAYDIN E. Multiple kernel learning algorithms[J]. *Journal of Machine Learning Research*, 2011, 12: 2211–2268.

[15] ÖZÖĞÜR-AKYÜZ S and WEBER G W. On numerical optimization theory of infinite kernel learning[J]. *Journal of Global Optimization*, 2010, 48(2): 215–239. doi: 10.1007/s10898-009-9488-x.

[16] 殷礼胜, 唐圣期, 李胜, 等. 基于整合移动平均自回归和遗传粒子群优化小波神经网络组合模型的交通流预测[J]. 电子与信息学报, 2019, 41(9): 2273–2279. doi: 10.11999/JEIT181073.
YIN Lisheng, TANG Shengqi, LI Sheng, *et al*. Traffic flow prediction based on hybrid model of auto-regressive integrated moving average and genetic particle swarm optimization wavelet neural network[J]. *Journal of Electronics & Information Technology*, 2019, 41(9): 2273–2279. doi: 10.11999/JEIT181073.

[17] 钱亚冠, 卢红波, 纪守领, 等. 基于粒子群优化的对抗样本生成算法[J]. 电子与信息学报, 2019, 41(7): 1658–1665. doi: 10.11999/JEIT180777.
QIAN Yaguan, LU Hongbo, JI shouling, *et al*. Adversarial example generation based on particle swarm optimization[J]. *Journal of Electronics & Information Technology*, 2019, 41(7): 1658–1665. doi: 10.11999/JEIT180777.

[18] https://opensky-network.org/.

王布宏：男，1975年生，博士，教授，研究方向为人工智能安全、信息物理系统安全等.

罗 鹏：男，1995年生，硕士生，研究方向为人工智能安全、ADS-B数据攻击检测.

李腾耀：男，1991年生，博士生，研究方向为ADS-B数据攻击检测和弹性恢复.

田继伟：男，1993年生，博士生，研究方向为人工智能安全、信息物理系统安全.

尚福特：男，1992年生，博士生，研究方向为信息物理系统安全.

责任编辑：余 蓉