ADS-B攻击数据弹性恢复方法

李腾耀[①] 王布宏*[①] 尚福特[①] 田继伟[①] 曹堃锐[①②]
[①](空军工程大学信息与导航学院 西安 710077)
[②](国防科技大学信息通信学院 西安 710106)

摘 要：为了对自动广播相关监视(ADS-B)攻击数据进行弹性恢复，确保空情态势感知信息的持续可用性，该文提出针对ADS-B攻击数据的弹性恢复方法。基于前置的攻击检测机制，获取当前ADS-B量测数据序列和预测数据序列，并在此基础上构建偏差数据序列、差分数据序列和邻近密度数据序列。依托偏差数据构建恢复向量，依托差分数据挖掘攻击数据的时序特性，依托邻近密度数据挖掘攻击数据的空间特性。通过整合3种数据序列构建弹性恢复策略并确定恢复终止点，实现对攻击影响的弱化，将ADS-B攻击数据向正常数据方向进行定向恢复。通过对6种典型攻击样式的实验分析，证明该弹性恢复方法能够有效恢复ADS-B攻击数据，削弱数据攻击对监视系统的影响。

关键词：空管监视；自动广播相关监视；数据安全；攻击检测；弹性恢复
中图分类号：TP915.08 **文献标识码**：A **文章编号**：1009-5896(2020)10-2365-09
DOI: 10.11999/JEIT191020

A Resilient Recovery Method on ADS-B Attack Data

LI Tengyao[①] WANG Buhong[①] SHANG Fute[①] TIAN Jiwei[①] CAO Kunrui[①②]
[①](*College of Information and Navigation, Air Force Engineering University, Xi'an* 710077, *China*)
[②](*School of Information and Communications, National University of Defense Technology, Xi'an* 710106, *China*)

Abstract: In order to conduct effective resilient recovery on Automatic Dependent Surveillance-Broadcast (ADS-B) attack data and ensure the continuous availability of air traffic situation awareness, a resilient recovery method on ADS-B attack data is proposed. Based on attack detection strategies, the measurement and prediction sequences of ADS-B data are obtained to set up deviation data, differential data and neighbor density data sequences, which are designed to construct recovery vectors, mine the temporal correlations and the spatial correlations respectively. The selected data sequences are integrated to accomplish the whole recovery method and decide the end point of recovery. The method is applied to elinimating attack effects and recovering the attack data towards normal data. According to the results of experiments on six classical attack patterns, the proposed method is effective on recovering attack data and eliminating the attack impacts.

Key words: Air traffic surveillance; Automatic Dependent Surveillance-Broadcast (ADS-B); Data security; Attack detection; Resilient recovery

1 引言

随着空域密度和飞行器种类的快速增长，当前空管系统已经难以满足其发展需求，正逐步向下一代网络化空管系统进行过渡，其智能化、自动化和一体化水平不断提升[1]。自动广播相关监视(Automatic Dependent Surveillance-Broadcast, ADS-B)技术作为下一代空管系统发展的核心技术之一，凭借高精度、广覆盖、易部署和支持空空监视等优势已经逐步发展成为下一代空管监视系统的主要监视手段[2]，为构建精准可靠的空情态势感知能力提供重要的数据基础[3]。但是，ADS-B协议在设计之初缺乏充分且完善的安全加固措施[4]。在传输过程中ADS-B报文以明文方式进行传输，缺乏必要的数据完整性验证手段和可靠的数据可认证性措施，使得ADS-B数据暴露在多样化的攻击威胁之下[5]。攻击者通过廉价的ADS-B接收设备能够对特定空域的监

收稿日期：2019-12-23；改回日期：2020-05-22；网络出版：2020-07-21
*通信作者： 王布宏 wbhgroup@aliyun.com
基金项目：国家自然科学基金(61902426)
Foundation Item: The National Natural Science Foundation of China (61902426)

视数据进行大规模采集，并在此基础上对ADS-B数据实施截获、篡改、伪造、重放和阻塞攻击[6]，使依托ADS-B数据构建的空情态势出现显著偏差，达到破坏监视系统可用性的目的。目前，针对ADS-B数据的攻击行为随着ADS-B部署范围的逐步扩大不断涌现，虚假数据注入攻击、幽灵飞行器注入攻击、飞行器数据泛洪攻击和数据重放攻击等攻击行为已经渗透到ADS-B数据中[7]，文献[8]在实验环境中验证了ADS-B数据攻击行为的可行性，文献[9]已经充分验证了针对ADS-B数据的攻击行为的有效性。

为了削弱攻击行为的影响，增强ADS-B数据的安全性，针对ADS-B数据的攻击检测机制成为防御策略的重要一环，能够将攻击数据屏蔽在决策流程之外[10]。虽然，针对ADS-B数据的攻击检测策略[11–14]能够快速准确地发现攻击数据和识别攻击行为，但是目前的研究尚缺乏有效且可靠的攻击响应机制。当攻击数据被识别后，ADS-B攻击数据序列会被简单地抛弃掉，在部分系统中会使用较低精度的其他监视数据源替代遭受攻击的数据序列，从而维持空管监视系统的可用性。但是由于不同监视数据源监视精度、覆盖范围、时间同步和数据关联等方面的差异性[15]，这样的处理机制不可避免地降低了监视系统的性能，在仅有ADS-B数据可供参考的空域或针对低延时突发事件处置时难以有效保证监视系统的可用性[16]。为了对攻击数据进行充分利用，实现攻击条件下ADS-B数据序列的有效重构，本文提出了ADS-B弹性恢复方法，在多样化攻击样式下削弱攻击行为对ADS-B数据的破坏性影响，保证ADS-B数据的持续可用性。在有效的攻击检测条件下，获取当前数据序列和对应的预测数据序列，据此分析构建偏差数据序列、差分数据序列和邻近密度数据序列，通过分析这些数据的概率分布和抖动幅值构建弹性恢复策略，提取攻击数据中包含的攻击向量，将攻击数据映射到正常数据的邻域范围，构建鲁棒可靠的空情态势感知能力。

2 问题建模

当ADS-B数据序列到达时，攻击检测机制将利用时空关联特性对数据进行分析并对攻击数据序列进行标记。假设攻击检测机制足够有效，攻击数据的起始采样点将被准确标记。设ADS-B数据序列为$\boldsymbol{x}_i$, ADS-B正常数据序列为$\boldsymbol{x}_{\rm n}$，攻击数据序列为$\boldsymbol{x}_{\rm a}$，攻击检测机制为$\boldsymbol{\Phi}$，攻击行为标记起始点为$a_{\rm s}$，则针对攻击检测策略可以建模为$\{\boldsymbol{x}_{\rm n},\boldsymbol{x}_{\rm a},a_{\rm s}\}=\boldsymbol{\Phi}(\boldsymbol{x}_i)$。通过对ADS-B数据进行攻击检测，能够获取攻击起始数据采样点$a_{\rm s}$，并将采集得到的数据序列$\boldsymbol{x}_i$以$a_{\rm s}$为边界划分为正常数据序列$\boldsymbol{x}_{\rm n}$和攻击数据序列$\boldsymbol{x}_{\rm a}$。对于攻击检测机制而言，如果采用生成式模型进行构造，会包含针对ADS-B流数据的预测器，用于实现对ADS-B数据的实时预测。该预测器一般用于描述ADS-B数据在正常模式下应当取的预测值，设该预测序列为$\boldsymbol{x}_{\rm p}$。预测序列$\boldsymbol{x}_{\rm p}$作为ADS-B数据序列$\boldsymbol{x}_i$在下一时间窗口的参考值，可以用于构建ADS-B数据的偏差序列，为攻击检测提供基本的判别依据。设预测器为ψ，则预测数据为$\boldsymbol{x}_{\rm p}=\psi(\boldsymbol{x}_i)$。在进行数据恢复的过程中，预测器是必要的。当攻击行为具有较强的不确定性和较高的动态性时，攻击数据恢复将采用预测数据作为填充数据，完成最低限度的弹性恢复功能。

如图1所示，针对ADS-B序列进行的弹性恢复是以可靠的攻击检测策略为前提的，并且攻击检测机制中需要包含有效的在线预测机制。当攻击检测完成并将当前数据标记为攻击数据后，可以确定攻击数据$\boldsymbol{x}_{\rm a}$和可靠的历史数据$\boldsymbol{x}_{\rm n}$。利用可靠的历史数据，能够得到重要的参考预测值。弹性恢复策略在攻击数据起始标记值$a_{\rm s}$处被触发，基于预测序列$\boldsymbol{x}_{\rm p}$和攻击数据序列$\boldsymbol{x}_{\rm a}$构建弹性恢复策略，确定攻击数据的恢复值以及恢复完成的终止点。设弹性恢复策略为Υ，恢复数据序列为$\boldsymbol{x}_{\rm r}$，恢复终止标记点为$a_{\rm e}$，则弹性恢复问题可以描述为：$\{\boldsymbol{x}_{\rm r},a_{\rm e}\}=\Upsilon(\boldsymbol{x}_{\rm a},\boldsymbol{x}_{\rm p},a_{\rm s})$。在恢复过程中，恢复数据序列将与参考数据序列进行偏差分析，确定恢复的终止标记点。恢复数据序列将持续替换攻击数据序列，向信息系统中的决策流进行推送。

3 弹性恢复方法

3.1 数据建模

考虑到信道噪声、量测误差等因素，包含噪声数据影响的ADS-B正常数据序列表示为$\boldsymbol{x}_{\rm n}=\boldsymbol{x}_{\rm o}+\boldsymbol{z}$，其中$\boldsymbol{x}_{\rm o}$为飞行器状态数据的真实值，$\boldsymbol{z}$为噪声数据。但是，噪声数据一般难以从正常数据中完全去除。针对正常数据构建的预测数据序列，作

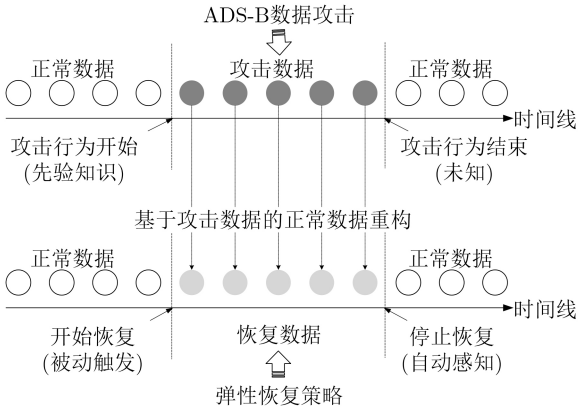


图 1 问题建模

为重要的参考数据建模为$\boldsymbol{x}_{\mathrm{p}}=\psi(\boldsymbol{x}_i)=\boldsymbol{x}_{\mathrm{n}}+\boldsymbol{x}_{\mathrm{m}}$，其中$\boldsymbol{x}_{\mathrm{m}}$为模型预测误差。预测数据必须与量测数据序列保持同步采样，当量测数据为正常数据序列时，预测数据与量测数据保持高度近似。对于攻击检测机制判定的攻击数据序列，需要重点关注其包含的攻击向量，因而将其建模为$\boldsymbol{x}_{\mathrm{a}}=\boldsymbol{x}_{\mathrm{n}}+\boldsymbol{a}_{\mathrm{v}}$，其中$\boldsymbol{a}_{\mathrm{v}}$为攻击向量。在基于攻击数据完成正常数据的重构恢复时，攻击向量必须得到削弱或者剔除，从而将其取值向正常数据序列$\boldsymbol{x}_{\mathrm{n}}$进行靠拢。

基于量测数据序列和预测数据序列构建偏差序列$\boldsymbol{x}_{\mathrm{d}}$，用以表征当前数据量偏离生成式模型的预测参考值的距离：$\boldsymbol{x}_{\mathrm{d}}^{(j)}=F_L(\boldsymbol{x}_{\mathrm{p}}^{(j)}-\boldsymbol{x}_i^{(j)})=\left|\boldsymbol{x}_{\mathrm{p}}^{(j)}-\boldsymbol{x}_i^{(j)}\right|$，其中$F_L$为距离度量函数。当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{n}}$时，$\boldsymbol{x}_{\mathrm{d}}^{(j)}=\left|\boldsymbol{x}_{\mathrm{m}}^{(j)}-\boldsymbol{z}^{(j)}\right|\to 0$；当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{a}}$时，$\boldsymbol{x}_{\mathrm{d}}^{(j)}=\left|\boldsymbol{x}_{\mathrm{m}}^{(j)}-\boldsymbol{z}^{(j)}-\boldsymbol{a}_{\mathrm{v}}^{(j)}\right|\sim\left|\boldsymbol{a}_{\mathrm{v}}^{(j)}\right|\gg 0$。因此，在攻击条件下，$\boldsymbol{x}_{\mathrm{d}}$携带了攻击向量的信息，可以作为数据恢复的参考基准。基于特定时间窗口内的量测数据构建差分序列$\boldsymbol{x}_{\mathrm{b}}$，用以表征当前数据量在指定时间窗口内相邻取值的变化幅值：$\boldsymbol{x}_{\mathrm{b}}^{(j)}=F_{\mathrm{b}}(\boldsymbol{x}_i^{(j+T_{\mathrm{w}})},\boldsymbol{x}_i^{(j)})=\left|\boldsymbol{x}_i^{(j+T_{\mathrm{w}})}-\boldsymbol{x}_i^{(j)}\right|$，其中$T_{\mathrm{w}}$为时间窗口长度，$F_{\mathrm{b}}$为差分数据度量函数。当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{n}}$时，$\boldsymbol{x}_{\mathrm{b}}^{(j)}=\left|\Delta\boldsymbol{x}_{\mathrm{n}}^{(j)}\right|\in[0,\mathrm{C_b}]$；当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{a}}$时，$\boldsymbol{x}_{\mathrm{b}}^{(j)}=\left|\Delta\boldsymbol{x}_{\mathrm{n}}^{(j)}+\Delta\boldsymbol{a}_{\mathrm{v}}^{(j)}\right|\gg\mathrm{C_b}$，其中$\mathrm{C_b}$为常量，一般取值为较小的判定阈值。在攻击条件下，差分数据能够描述攻击行为导致的时序关联特性的破坏量，为恢复数据的时序关联特性提供重要支持。根据量测到的不同目标的状态数据序列，分析当前飞行器的邻近密度数据序列$\boldsymbol{x}_{\mathrm{g}}$，用以表征当前飞行器在特定采样时间周期内对应空域飞行器的密度变化情况：$\boldsymbol{x}_{\mathrm{g}}=N(\boldsymbol{x}_i)/F_{\mathrm{g}}(\boldsymbol{x}_i)$，其中$F_{\mathrm{g}}$表示以当前飞行器为球心测算的空间范围，$N$表示当前飞行器在$F_{\mathrm{g}}$内邻近飞行器的数量。当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{n}}$时，$\left|\Delta\boldsymbol{x}_{\mathrm{g}}^{(j)}\right|\to\delta_{\mathrm{o}}$，其中$\delta_{\mathrm{o}}$为0的邻域；当$\boldsymbol{x}_i=\boldsymbol{x}_{\mathrm{a}}$时，$\left|\Delta\boldsymbol{x}_{\mathrm{g}}^{(j)}\right|>C_{\mathrm{g}}$，其中$C_{\mathrm{g}}$为浮动的判定阈值。邻近密度数据序列充分利用了ADS-B数据的空间关联特性，当攻击行为注入大量飞行器目标时，$\Delta\boldsymbol{x}_{\mathrm{g}}$会出现较大波动。

3.2 恢复方法

在针对ADS-B攻击数据进行恢复时，通过综合考虑该数据序列在攻击影响下的变化、在时空关联下的不确定性波动，使用偏差数据$\boldsymbol{x}_{\mathrm{d}}$、差分数据$\boldsymbol{x}_{\mathrm{b}}$和邻近密度数据$\boldsymbol{x}_{\mathrm{g}}$构建描述攻击影响的度量集合$\boldsymbol{X}=\{\boldsymbol{x}_{\mathrm{d}},\boldsymbol{x}_{\mathrm{b}},\boldsymbol{x}_{\mathrm{g}}\}$。通过分析度量集合$\boldsymbol{X}$中序列变化趋势，定制攻击数据恢复方法，实现攻击数据向正常数据的有效映射。

3.2.1 偏差数据分析

当获取到偏差数据序列$\boldsymbol{x}_{\mathrm{d}}$时，分析其变化趋势。首先，利用线性模型进行估计，以便应对常量偏差注入攻击引入的破坏效果，设当前偏差数据序列满足线性模型$f(\boldsymbol{x}_{\mathrm{d}})=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}_{\mathrm{d}}+\boldsymbol{b}$，考虑到模型误差和噪声影响，使用支持向量回归(Support Vector Regression, SVR)对偏差数据的潜在规律进行估计，利用拉格朗日乘子法获得该线性模型的解集合$\{\boldsymbol{\omega},\boldsymbol{b}\}$。当线性模型拟合误差的期望值在判定时间窗口内持续低于阈值δ时，接受线性模型的假设，利用估计量$\hat{f}(\boldsymbol{x}_{\mathrm{d}})$对攻击数据构建恢复量$\boldsymbol{\chi}$，作为攻击数据的替代值：

$$\boldsymbol{\chi}=\boldsymbol{x}_{\mathrm{a}}-\hat{f}(\boldsymbol{x}_{\mathrm{d}}) \tag{1}$$

当偏差数据序列无法被线性模型拟合时，利用混合高斯模型(Gaussian Mixture Model, GMM)对偏差数据序列$\boldsymbol{x}_{\mathrm{d}}$满足的概率分布进行估计，以应对随机偏差注入攻击对ADS-B序列的影响。利用时间窗口内的偏差数据序列作为GMM模型的输入，对不同数据点满足的高斯分布参数进行估计，获得描述攻击数据偏差的混合高斯模型。进而，针对不同偏差数据点$\boldsymbol{x}_{\mathrm{d}}{}^{(i)}$，利用其所对应的高斯分布获取期望值$\mathrm{E}(\boldsymbol{x}_{\mathrm{p}}{}^{(i)})$并构建恢复量$\boldsymbol{\chi}$

$$\boldsymbol{\chi}=\boldsymbol{x}_{\mathrm{a}}{}^{(i)}+(-1)^{(\hat{f}(\boldsymbol{x}_{\mathrm{d}})^{(i)}-\boldsymbol{x}_{\mathrm{d}}^{(i)})}\mathrm{E}(\boldsymbol{x}_{\mathrm{p}}{}^{(i)}) \tag{2}$$

3.2.2 差分数据分析

差分数据是对量测数据序列时序关联的度量，对数据篡改攻击较为敏感，能够对数据重放攻击进行针对性恢复。针对相邻量测数据值构建基本的差分数据序列$\boldsymbol{x}_{\mathrm{b}}$，当$\boldsymbol{x}_{\mathrm{b}}$出现突变时，设突变数据点为$i=k$，有$\boldsymbol{x}_{\mathrm{b}}{}^{(k)}>\mathrm{C_b}$。对突变数据点后续数据序列，在时间长度$l_t$内分析$\boldsymbol{x}_{\mathrm{b}}$的幅值变化。设定差分数据变化的可信阈值为$C_\tau$，当$\Delta\boldsymbol{x}_{\mathrm{b}}^{(i)}>C_\tau$时，说明当前数据已经被恶意篡改，并且攻击者没有按照飞行器目标的状态连续变化构造攻击向量。为了应对该情况下的攻击行为，使用预测数据序列对攻击数据进行填充替换。当$\Delta\boldsymbol{x}_{\mathrm{b}}^{(i)}\le C_\tau$时，说明当前数据没有超出飞行器目标状态连续变化的基本约束，位置和速度信息仍然满足基本的运动规律。在飞行器状态数据重置和数据重放攻击条件下，易出现该类ADS-B量测序列。为了有效恢复相邻数据的连续变化，构建时间窗口集$T=\{t_i|t_i=\gamma_i t_{\mathrm{o}},\gamma_i\in N_+\}$，其中$t_{\mathrm{o}}$是两组ADS-B数据的发送时间间隔。通过对$\gamma_i$进行有效选择，在不同时间窗口条件下对数据的周期性变化规律进行挖掘。为了实现遍历效率和分析性能的权衡，令$\gamma_i=2^i,i\in N_+$，其中将γ_i选取值的数量上限为n_γ，即$|\{\gamma_i\}|<n_\gamma$。进一步地，对特定时间周期内数据的变化进行度量。

$$\gamma_{\mathrm{c}}=\arg\min_{\gamma_i}\left(\sum_{j\in l_t}\left|\boldsymbol{x}_{\mathrm{a}}{}^{(j)}-\boldsymbol{x}_{\mathrm{a}}{}^{(j-\gamma_i t_{\mathrm{o}})}\right|\right) \tag{3}$$

当该度量值的变化均值稳定在较小范围内时，攻击数据潜在的攻击模式可能为重放攻击行为

$$\varepsilon=\frac{1}{l_t}\sum_{j\in l_t}\left|\boldsymbol{x}_{\mathrm{a}}{}^{(j)}-\boldsymbol{x}_{\mathrm{a}}{}^{(j-\gamma_{\mathrm{c}}t_{\mathrm{o}})}\right|\le\phi C_\tau,\phi\in[0,1] \tag{4}$$

其中，ε为评估数据周期性变化的关键指标，当数据在时间维度产生平移后，ε会处在较小的取值空间中。特别地，当γ_{c}与攻击者实施重放攻击的时间差相匹配时，$\varepsilon\to 0$。当出现如式(4)所示的情况时，将当前攻击数据序列在时间维度向前平移γ_{c}个时间长度，并利用预测数据序列对攻击数据最后γ_{c}个数据值进行预测填充。当$\varepsilon>\phi C_\tau$时，依托预测器对攻击窗口内的数据进行预测，并利用偏差数据序列进行定向恢复。

3.2.3 邻近密度数据分析

邻近密度数据是在空域飞行密度快速增长的过程中度量目标飞行器状态数据空间关联性的重要指标。以当前被攻击目标为球心构建邻近密度测算范围K，设ADS-B接收器接收半径为r_{c}，选定的邻近密度测算半径为$R_{\mathrm{d}}=\kappa r_{\mathrm{c}},\kappa\in[0,1]$。在时间窗口$T_{\mathrm{d}}$内，如果遭受攻击的飞行器目标邻近密度数据出现快速增长时，即$\Delta\boldsymbol{x}_{\mathrm{g}}>C_{\mathrm{g}}$，此时攻击数据可能受到飞行器泛洪攻击影响。一般地，C_{g}取值会高于正常状态下邻近密度取值的最大值。

$$\frac{\tau}{T_{\mathrm{d}}}\sum_{j\in T_{\mathrm{d}}}\left|\Delta\boldsymbol{x}_{\mathrm{g}}^{(j)}\right|\ge C_{\mathrm{g}} \tag{5}$$

其中，τ为ADS-B消息的接收时间间隔，当ADS-B数据在时间窗口T_{d}内持续出现邻近密度快速增长时，当前飞行器为幽灵飞行器的概率随之增长。设在邻近范围K内，遭受攻击的目标飞行器数量为n_{a}，对应的全部目标数量为n_{w}，则定义攻击目标密度为

$$\rho=\frac{1}{\max(n_{\mathrm{a}})}\sum_j\frac{n_{\mathrm{a}}{}^{(j)}}{n_{\mathrm{w}}{}^{(j)}},j\in K \tag{6}$$

如果在当前遭受攻击的飞行器目标邻近范围内，出现多个攻击数据序列，并且均满足$\rho>C_\rho$，则将标记为攻击数据序列的目标进行抛弃。如果当前范围内仅存在1个攻击数据序列，则借助差分数据对当前数据序列进行恢复。

3.2.4 弹性恢复策略

通过对偏差数据序列、差分数据序列和邻近密度数据序列等数据的综合分析，针对单个攻击数据序列构建弹性恢复策略。对ADS-B数据进行实时采集以获取量测序列$\boldsymbol{x}_i$，根据系统已经部署的攻击检测算法获取ADS-B预测序列$\boldsymbol{x}_{\mathrm{p}}$和攻击初始点$a_{\mathrm{s}}$。根据$a_{\mathrm{s}}$将量测序列$\boldsymbol{x}_i$划分为正常数据序列$\boldsymbol{x}_{\mathrm{n}}$和攻击数据序列$\boldsymbol{x}_{\mathrm{a}}$，构建偏差数据序列$\boldsymbol{x}_{\mathrm{d}}$。如果$\boldsymbol{x}_{\mathrm{d}}$满足线性变化规律，直接对攻击向量进行提取，利用SVR方法构建恢复向量$\boldsymbol{\chi}$；如果$\boldsymbol{x}_{\mathrm{d}}$满足非线性规律，利用GMM方法对其潜在的分布规律进行分析，构建恢复向量$\boldsymbol{\chi}$作为后续分析的基础。根据$\boldsymbol{x}_{\mathrm{a}}$构建差分数据序列$\boldsymbol{x}_{\mathrm{b}}$，如果差分数据序列在某个特定的时间点超出可信阈值$\mathrm{C_b}$，则判定当前遭受时序攻击的概率较高。如果破坏效果在$\Delta\boldsymbol{x}_{\mathrm{b}}$中持续存在，则使用恢复向量$\boldsymbol{\chi}$进行填充；如果破坏效果在$\Delta\boldsymbol{x}_{\mathrm{b}}$序列中仅在有限时间点处出现突变，在时间窗口集$T$中进行遍历，选择使得对应时间长度下差分数据累积和最小的时间窗口标记值γ_{c}，并在该时间窗口$\gamma_{\mathrm{c}}t_{\mathrm{o}}$下分析数据周期性变化指标$\varepsilon$。如果$\varepsilon\le\phi C_\tau$，则在时间维度上将$x_{\mathrm{a}}$向前平移$\gamma_{\mathrm{c}}$个时间长度并利用$\boldsymbol{x}_{\mathrm{p}}$对攻击数据$\gamma_{\mathrm{c}}$个数据值进行预测填充，否则利用预测数据序列$\boldsymbol{x}_{\mathrm{p}}$对整个攻击数据序列$\boldsymbol{x}_{\mathrm{a}}$进行预测填充。进一步地，利用量测数据序列$\boldsymbol{x}_i$分析邻近密度数据序列$\boldsymbol{x}_{\mathrm{g}}$，当邻近密度数据序列的变化量$\Delta\boldsymbol{x}_{\mathrm{g}}$的统计量超出可信阈值$C_{\mathrm{g}}$且攻击目标密度值$\rho$超出控制上限$C_\rho$时，对应攻击数据序列将以大概率认定为幽灵飞行器目标，因而将该量测数据序列进行抛弃，并进行标记。

经过上述综合分析后，能够得到有效的弹性恢复向量$\boldsymbol{x}_{\mathrm{r}}$，并在此基础上对攻击数据序列进行恢复。基于恢复数据序列$\boldsymbol{x}_{\mathrm{r}}$与量测数据序列$\boldsymbol{x}_i$的偏差构建终止点$a_{\mathrm{e}}$的确定规则。设定时间维度上的滑动窗口并构建数据采样时间点集合T_{r}，其采样间隔等同于相邻ADS-B数据报文接收的平均时间间隔τ，滑动窗口长度为$|T_{\mathrm{r}}|$。在时间采样点集合T_{r}内，当恢复数据序列和量测数据序列持续近似时，ADS-B数据逐步达到弹性恢复的终止点a_{e}。

$$\left.\begin{aligned}&a_{\mathrm{e}}=\min(T_{\mathrm{r}}),\\&\text{s.t.}\quad\frac{1}{|T_{\mathrm{r}}|}\sum_{i\in T_{\mathrm{r}}}\left|\boldsymbol{x}_{\mathrm{r}}{}^{(i)}-\boldsymbol{x}_i{}^{(i)}\right|\to 0\end{aligned}\right\} \tag{7}$$

一旦a_{e}被确定，恢复过程将被终止，并将采集得到的量测数据流作为正常数据送入空管监视系统的后续处理流程。

3.3 策略评估

当弹性恢复策略应用到攻击数据序列时，攻击行为造成的破坏影响将得到一定程度的削弱。为了对该恢复策略进行有效评估，使用第三方参考监视数据源作为正常可靠数据对恢复数据进行比对分析。在进行评估的过程中，会使用到多种监视数据

源数据，这是为了防止单个数据源引入的噪声和攻击不确定性。因为攻击者对所有监视手段同时实施攻击的可能性比较小，因而使用除ADS-B数据之外的多种参考数据源构建时间同步的正常数据序列$\boldsymbol{x}_{\mathrm{f}}$。基于参考数据源$\boldsymbol{x}_{\mathrm{f}}$、攻击数据$\boldsymbol{x}_{\mathrm{a}}$和恢复数据序列$\boldsymbol{x}_{\mathrm{r}}$构建评估恢复策略有效性的指标$Q$

$$Q=\ln \mathrm{E}\left(\left|\frac{x_{\mathrm{a}}-x_{\mathrm{r}}}{x_{\mathrm{f}}-x_{\mathrm{r}}+\eta}\right|\right)+\varphi\frac{\mathrm{D}(x_{\mathrm{f}}-x_{\mathrm{r}})}{\mathrm{D}(x_{\mathrm{a}}-x_{\mathrm{r}})} \tag{8}$$

其中，E为期望函数，D为方差函数，φ为权重常量，η为避免分母为0的微小修正量。评估指标Q重点考虑了恢复策略对攻击影响的削弱程度以及恢复过程中的稳定性。当恢复序列逼近正常取值并且稳定地分布在正常值邻近范围时，Q的取值会增大；当恢复序列远离正常取值或者恢复序列仅仅恢复小比例攻击数据时，Q的取值会减小。

4　实验分析

在实验过程中，针对6种常见的攻击样式应用弹性恢复方法，对偏差数据序列、差分数据序列和邻近密度数据序列的特性进行分析，实现对弹性恢复策略有效性的验证。本实验以真实ADS-B数据集作为分析基础，所有数据获取自OpenSky项目[17]，结合攻击行为对数据序列的攻击影响构造实验分析所需的攻击数据集。以文献[14]提出的攻击检测方法作为基础，利用层次时序记忆模型构建ADS-B量测序列的在线预测数据序列，随机选取攻击目标并应用本文提出的弹性恢复方法完成分析验证。

4.1　攻击行为分析

由于安全法规的限制，在真实空域环境中ADS-B攻击数据难以有效采集。通过对ADS-B数据攻击模式的分析，依托表1所示的典型攻击样式构造特定攻击数据序列。对于常量偏差注入攻击(ATK-1)，攻击向量构造为$\boldsymbol{a}_{\mathrm{v}}=\mathcal{K}\boldsymbol{x}_{\mathrm{n}},\mathcal{K}\in[0,0.1]$，偏差注入比例$\mathcal{K}$控制在对应属性数据取值的10%以内，在正常数据序列中附加常量偏差，实现对正常数据的定向篡改；对于随机偏差注入攻击(ATK-2)，攻击向量构造为$\boldsymbol{a}_{\mathrm{v}}=\lambda\mathcal{K}_0\boldsymbol{x}_{\mathrm{n}}$，偏差注入初始比例$\mathcal{K}_0$略大于$\mathcal{K}$，随机数值因子$\lambda\sim N(0,1)$，实现在正常数据序列中附加一定比例的随机值，使攻击数据在正常数据邻近范围抖动；对于增量偏差注入攻击，攻击向量构造为$\boldsymbol{a}_{\mathrm{v}}^{(i)}=\widetilde{\mathcal{K}}_0\boldsymbol{x}_{\mathrm{n}}^{(i)}(t_i-t_0)/\tau$，偏差注入基础比例$\widetilde{\mathcal{K}}_0$略小于$\mathcal{K}$，$(t_i-t_0)/\tau$为当前数据采样点距离攻击初始点$a_s$的时间间隔，随着时间间隔的逐渐增长，对应的注入偏差取值按比例递增，使正常数据的偏差随时间不断扩大；对于航迹替换攻击，攻击向量构造为$\boldsymbol{a}_{\mathrm{v}}=-\boldsymbol{x}_{\mathrm{n}}+\boldsymbol{x}'_{\mathrm{n}}$, $\boldsymbol{x}'_{\mathrm{n}}$为替换数据，设定为邻近飞行器在等长时间窗口的数据序列，是正常数据在攻击窗口内被直接替换；对于航迹重放攻击，攻击向量构造为$\boldsymbol{a}_{\mathrm{v}}^{(i)}=-\boldsymbol{x}_{\mathrm{n}}^{(i)}+\boldsymbol{x}_{\mathrm{n}}^{(i-\mathcal{T})}$，其中时延控制参数$\mathcal{T}>5\tau$，将时间间隔$\mathcal{T}$之前的数据替换当前时间点的数据，间接造成正常数据的接收时延增大。对于飞行器泛洪攻击，是向攻击空域注入大量虚假飞行器数据，虚假飞行器数据通过随机构造飞行器的位置信息构成，注入的数量在30架次到300架次之间。利用攻击检测算法构建在线预测序列并对攻击数据进行标记，构建攻击影响度量集X，作为弹性恢复策略的输入。

4.2　数据序列分析

针对偏差数据、差分数据和邻近密度数据序列，结合6种攻击模式进行分析获取。ATK-1～ATK-33种攻击模式适用于偏差数据分析，并利用恢复向量进行恢复；ATK-4～ATK-5两种攻击模式适用于利用差分数据进行分析，对ATK-5进行时间平移能够有效削弱攻击影响；ATK-6攻击模式适合使用邻近密度数据序列进行恢复。在弹性恢复策略中，对于未知数据会首先根据偏差数据计算恢复向量，而后基于差分数据进行数据恢复，当攻击数据不适合通过时间平移进行恢复时会使用恢复向量进行数据替代，最后基于邻近密度数据针对幽灵飞行器注入攻击进行防御，及时剔除虚假的飞行器节点。

4.2.1　偏差数据分析

偏差数据是对攻击数据偏离正常数据潜在合规模式的有效度量。如图2所示，常量偏差注入攻击偏差数据和增量偏差注入攻击偏差数据可以使用线

表 1　构造的典型攻击样式

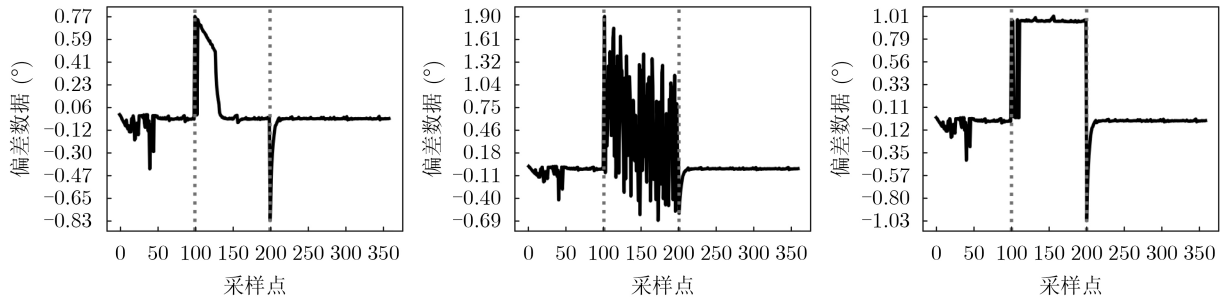
编号	攻击模式	攻击影响
ATK-1	常量偏差注入攻击	针对ADS-B多属性数据注入常量偏差
ATK-2	随机偏差注入攻击	针对ADS-B多属性数据注入随机偏差
ATK-3	增量偏差注入攻击	针对ADS-B多属性数据注入增量偏差
ATK-4	航迹替换攻击	针对特定时间窗口内的航迹进行替换
ATK-5	航迹重放攻击	在特定时间长度下实现航迹重放
ATK-6	飞行器泛洪攻击	向当前空域态势中注入大量幽灵飞行器目标

性模型进行拟合；随机偏差注入攻击偏差数据出现随机波动，使用概率分布估计能够对该数据序列进行描述。航迹替换攻击偏差数据频繁出现跳变，航迹重放攻击偏差数据在攻击开始和结束时出现波动，这两种攻击模式不易于直接使用基于偏差数据的恢复策略进行恢复。对常量偏差注入攻击偏差数据和增量偏差注入攻击偏差数据使用SVR进行拟合，能够对产生的偏差进行更好地估计，在下一时刻数据遭受破坏之前对偏差进行提前感知。对随机偏差注入攻击偏差数据使用GMM进行均值的分析，能够使用该均值对攻击数据进行恢复。

如图3所示，对ATK-1～ATK-3 3种攻击模式应用恢复策略，可以发现在攻击采样点100到攻击采样点200之间，恢复数据更加靠近正常数据附近。虽然仍旧保留了一定的攻击特征，但是在数值上显著地削弱了不同攻击模式的影响。

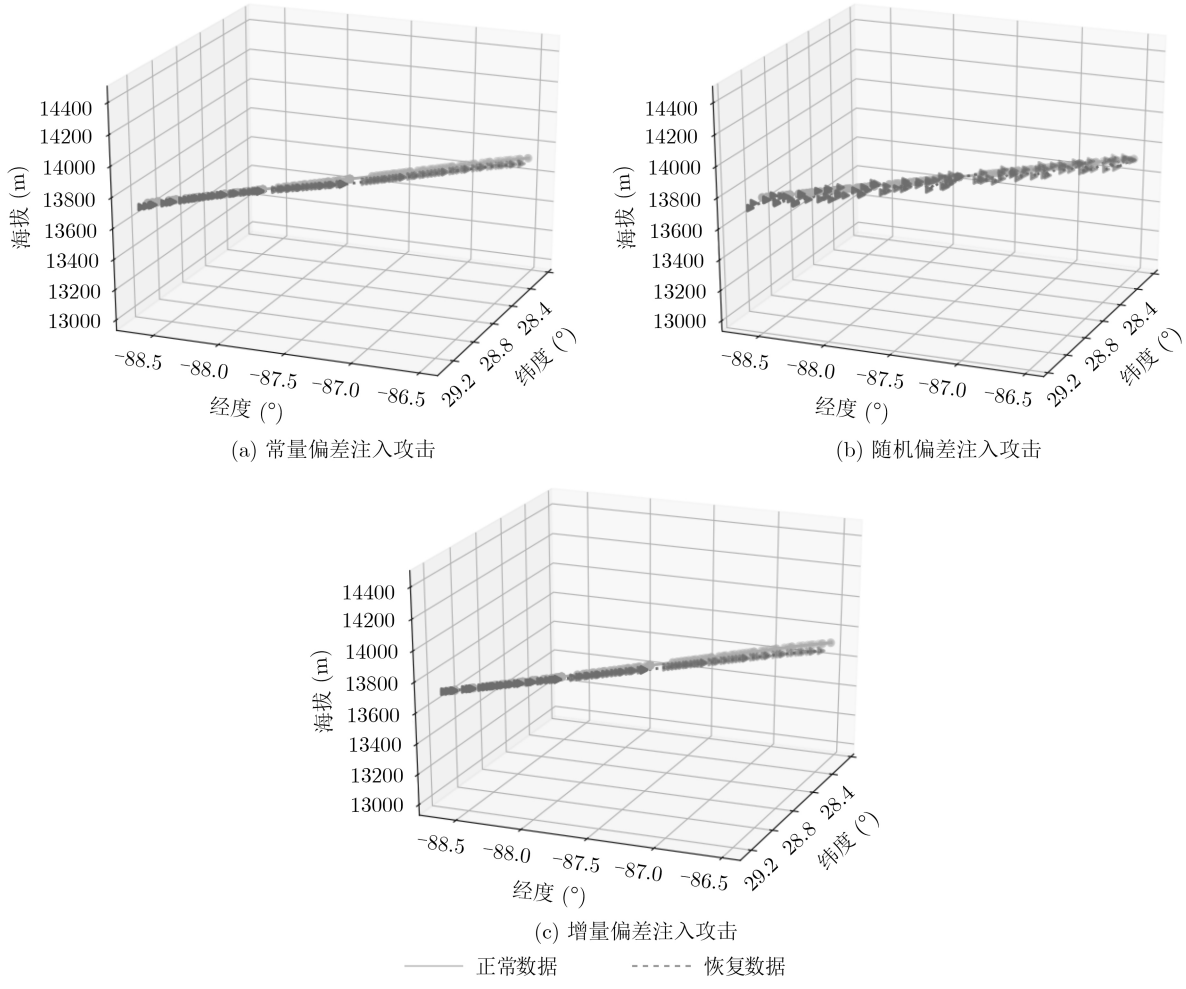
4.2.2 差分数据分析

差分数据用于分析ADS-B数据序列的时序关联



(a) 常量偏差注入攻击 (b) 随机偏差注入攻击 (c) 增量偏差注入攻击

图 2 偏差数据分析



(a) 常量偏差注入攻击 (b) 随机偏差注入攻击

(c) 增量偏差注入攻击

图 3 数据恢复效果(ATK-1～ATK-3)

特性，如图4所示。航迹替换攻击的差分数据在攻击起始点和终止点出现突变；航迹重放攻击的差分数据同样在起始点和终止点出现显著变化，但是幅度相较于航迹替换攻击要小很多，在攻击过程中在较小范围内出现浮动。

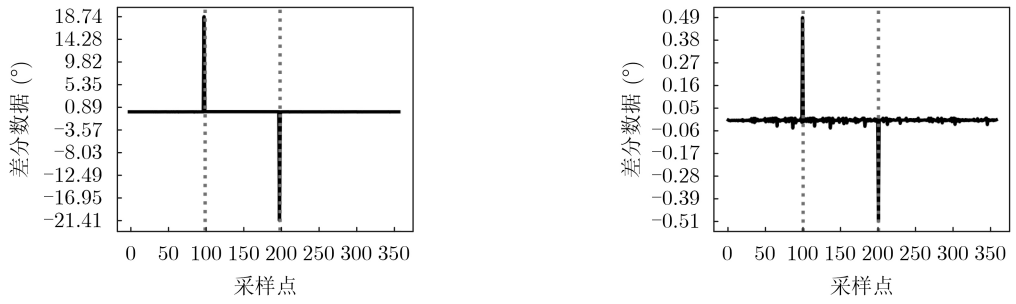
通过对航迹替换攻击和航迹重放攻击两种攻击模式进行恢复，攻击数据在一定程度上得到纠正，如图5所示。航迹替换攻击数据的恢复使用预测数据进行替换，由于替换航迹的不确定性，其恢复过程很不稳定，并仍旧与原航迹保持一定偏差；航迹重放攻击数据的恢复通过平移时间的最优化完成数据的逆向恢复，使攻击引入的时延得到修正。

4.2.3 邻近密度数据分析

通过对不同幽灵飞行器注入规模下的邻近密度数据进行采集分析，可以发现在该攻击模式下飞行器的邻近密度会出现快速增长，为进一步判别和标定飞行器泛洪攻击数据及攻击目标密度提供了支撑。如图6所示，通过应用恢复策略能够对幽灵飞行器进行删除，使空情态势遭受的破坏程度得到削弱，但是由于攻击目标密度在与真实飞行器较邻近范围内比较小的原因，存在部分幽灵飞行器没有被完全剔除。这使得针对飞行器泛洪攻击的数据恢复存在一定的限制。

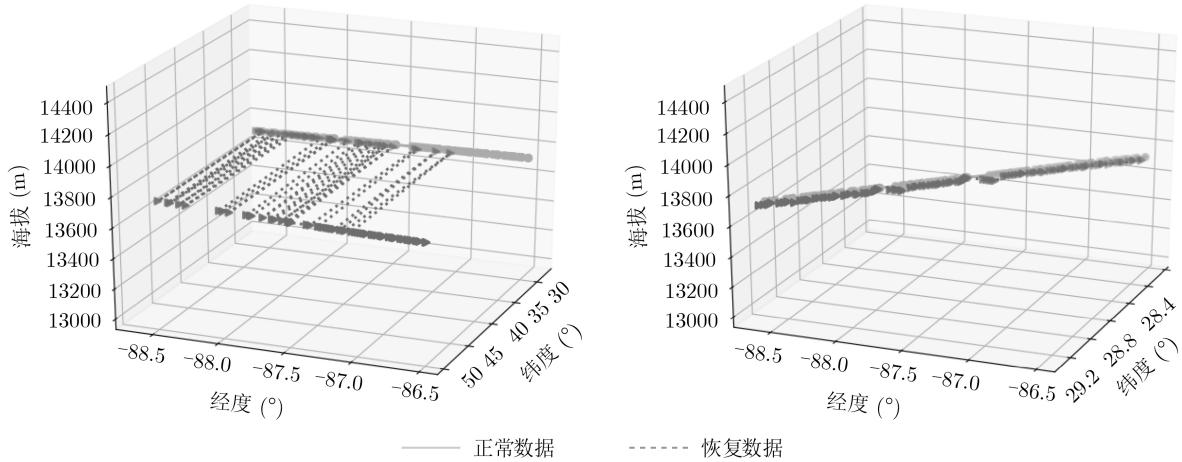
4.3 恢复策略分析

通过对6种典型攻击模式进行弹性恢复，能够获取到恢复数据和正常数据之间的差异数据。如图7所示，通过对Q值的分析，能够获取恢复策略对不同攻击模式的恢复效能。恢复策略对ATK-1～ATK-3的恢复效果更好，偏差注入攻击在有效的偏差估计方法下能够得到有效削弱。对常量偏差注入攻击(ATK-1)数据的恢复相对较差，这主要是由于常量偏差仅在攻击起始点位置出现大幅度波动，同时较小的常量偏差数值难以从攻击数据序列中有效提取；航迹替换攻击(ATK-4)由于替换航迹的不确定性，其恢复性能在6种典型攻击数据中是最差的；航迹重放攻击(ATK-5)受限于重放时延的搜索，恢复性能受到一定影响；飞行器泛洪攻击数据在恢复策略下能够对虚假飞行器目标进行一定程度上的剔除，部分与合法飞行器距离较近的飞行器数据不易被及时滤除，空情态势感知的可用性只能在一定程度上得到恢复。



(a) 航迹替换攻击　　(b) 航迹重放攻击

图 4 差分数据分析



(a) 航迹替换注入攻击　　(b) 航迹重放攻击

图 5 数据恢复效果(ATK-4, ATK-5)

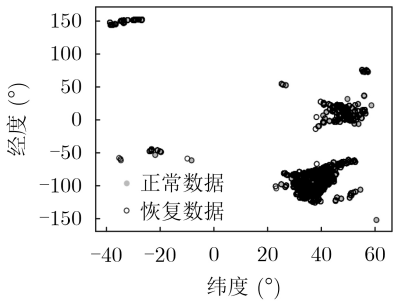


图 6 数据恢复效果(ATK-6)

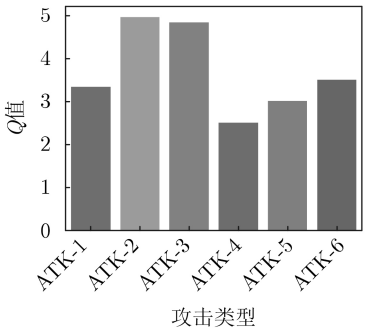


图 7 恢复策略效能分析

5 结束语

为了弥补ADS-B数据攻击检测策略缺乏有效数据恢复机制的不足，充分利用ADS-B攻击数据携带的合规数据信息，提出针对ADS-B数据的弹性恢复策略。对ADS-B攻击数据序列进行正常状态信息的重构，利用偏差数据序列、差分数据序列和邻近密度数据序列构建弹性恢复策略，实现对多样化攻击模式的有效恢复，确保空情态势感知的可用性。

但是，提出的恢复策略中存在多个可信阈值需要进行设定，这些阈值对数据恢复性能具有重要影响，在实践过程中主要利用可靠的历史数据的统计信息进行估计，但是由于ADS-B数据自身存在的动态性使得阈值的有效确定仍是恢复策略中需要持续改进和完善的重要方面。此外，前置攻击检测算法中的预测数据的可靠性对ADS-B数据弹性恢复方法的性能也会带来一定的影响，在进行数据恢复过程中仍需要进一步提升对攻击行为识别的准确性，以制定更具针对性的弹性恢复策略。

参 考 文 献

[1] STROHMEIER M, SCHÄFER M, PINHEIRO R, *et al*. On perception and reality in wireless air traffic communication security[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(6): 1338–1357. doi: 10.1109/tits.2016.2612584.

[2] WANG Wenyi, WU Renbiao, and LIANG Junli. ADS-B signal separation based on blind adaptive beamforming[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(7): 6547–6556. doi: 10.1109/TVT.2019.2914233.

[3] SUN Junzi, VÛ H, ELLERBROEK J, *et al*. pyModeS: Decoding mode-S surveillance data for open air transportation research[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 21(7): 2777–2786. doi: 10.1109/TITS.2019.2914770.

[4] STROHMEIER M, SCHÄFER 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.

[5] 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.

[6] MANESH M R and KAABOUCH N. Analysis of vulnerabilities, attacks, countermeasures and overall risk of the automatic dependent surveillance-broadcast (ADS-B) system[J]. *International Journal of Critical Infrastructure Protection*, 2017, 19: 16–31. doi: 10.1016/j.ijcip.2017.10.002.

[7] 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.

[8] SCHÄFER M, LENDERS V, and MARTINOVIC I. Experimental analysis of attacks on next generation air traffic communication[C]. The 11th International Conference on Applied Cryptography and Network Security, Berlin, Germany, 2013: 253–271. doi: 10.1007/978-3-642-38980-1_16.

[9] COSTIN A and FRANCILLON A. Ghost in the air (traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices[C]. Black Hat, Las Vegas, USA, 2012: 1–10.

[10] 陈红松, 陈京九. 基于循环神经网络的无线网络入侵检测分类模型构建与优化研究[J]. 电子与信息学报, 2019, 41(6): 1427–1433. doi: 10.11999/JEIT180691.

CHEN Hongsong and CHEN Jingjiu. Recurrent neural networks based wireless network intrusion detection and classification model construction and optimization[J]. *Journal of Electronics & Information Technology*, 2019, 41(6): 1427–1433. doi: 10.11999/JEIT180691.

[11] YING Xuhang, MAZER J, BERNIERI G, *et al*. Detecting ADS-B spoofing attacks using deep neural networks[C]. 2019 IEEE Conference on Communications and Network Security, Washington, USA, 2019: 187–195. doi: 10.1109/

CNS.2019.8802732.

[12] 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.

[13] 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.

[14] LI Tengyao, WANG Buhong, SHANG Fute, *et al*. Online sequential attack detection for ADS-B data based on hierarchical temporal memory[J]. *Computers & Security*, 2019, 87: 101599. doi: 10.1016/j.cose.2019.101599.

[15] ZHANG Tao, WU Renbiao, LAI Ran, *et al*. Probability hypothesis density filter for radar systematic bias estimation aided by ADS-B[J]. *Signal Processing*, 2016, 120: 280–287. doi: 10.1016/j.sigpro.2015.09.012.

[16] SMITH M, STROHMEIER M, HARMAN J, *et al*. Safety vs. security: Attacking avionic systems with humans in the loop[J]. arXiv, 2019, 1905.08039.

[17] STROHMEIER M, MARTINOVIC I, FUCHS M, *et al*. Opensky: A swiss army knife for air traffic security research[C]. The 34th IEEE/AIAA Digital Avionics Systems Conference, Prague, Czech Republic, 2015: 1–14. doi: 10.1109/DASC.2015.7311411.

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

王布宏：男，1975年生，博士，教授，研究方向为空管安全、信息物理系统安全和人工智能安全.

尚福特：男，1992年生，博士生，研究方向为信息物理系统安全.

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

曹堃锐：男，1989年生，博士生，研究方向为无线网络物理层安全.

责任编辑：陈　倩