# 自适应增强的动态网络流量主动异常检测

李彬[1]，王意洁[1]，程力[2]
[1] 国防科技大学计算机学院并行与分布计算全国重点实验室，中国长沙市，410073
[2]国防科技大学系统工程学院，中国长沙市，410073

**摘要：**主动异常检测通过查询被采样实例的标签，增量更新检测模型，已被广泛用于检测网络攻击。然而，现有方法不能在动态网络流量上实现预期表现，这是因为：（1）它们的查询策略不能采样具有信息量的网络流量，以使检测模型适应数据分布不断变化的网络流量；（2）它们的模型更新仅依赖于有限的查询流量，不能利用网络流量中巨大的未标记流量。为解决这些问题，提出一种自适应增强的主动先验知识森林模型$A^3$PF，用于网络流量的异常检测。通过利用网络攻击的先验知识，寻找能更好区分异常网络流量和正常网络流量的特征子空间，从而构建先验知识森林模型。一方面，为使模型适应不断变化的网络流量，设计了一种新的自适应查询策略，从动态数据分布的变化和异常的不确定性两个方面对具有信息量的网络流量进行采样。另一方面，基于邻域中网络流量的相似性，设计了一种增强更新方法，为查询流量的未标记邻居生成伪标签，从而在异常检测模型更新过程中能够充分利用大量未标记流量。在CIC-IDS2017和UNSW-NB15这两个入侵检测数据集上的大量实验表明，较之相关方法，$A^3$PF性能显著提升。具体而言，其平均AUC-ROC分别提高20.9%和21.5%，平均AUC-PR分别提高44.6%和64.1%。

**关键词：**主动异常检测；网络流量；伪标签；网络攻击的先验知识
https://doi.org/10.1631/FITEE.2300244