# 基于深度森林的网络匿名流量检测方法研究与应用

李成豪，\*\*\*，\*\*\*

（南京理工大学，江苏 南京 450001）

表5.4 对比实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 分类器 | 准确率% | 精确率% | F1-Score% | 召回率% |
| gcCS | 99.86 | 99.82 | 99.72 | 99.62 |
| gc | 99.87 | 99.82 | 99.68 | 99.75 |
| SVM | 98.43 | 97.00 | 94.66 | 92.43 |
| RandomForest | 99.79 | 99.64 | 99.31 | 98.98 |
| DecisionTree | 99.69 | 99.07 | 98.99 | 98.9 |
| DNN | 98.49 | 97.47 | 94.85 | 92.37 |

由表5.4可知,原始的深度森林模型和改进的深度森林模型取得的分类效果都略高其他方法。综合表5.3和表5.4可知改进的深度森林模型无论较原始深度森林的准确率和时间空间开销还是和其他方法的测试精度比较都取得了良好的效果。

## **6** 总结与展望

本文主要进行了Tor匿名通信系统产生的网络流量的检测与识别。首先通过分析Tor匿名流量与正常普通流量的特征。通过特征选择算法，本文提出了一组基于时间特征来检测并描述Tor流量。Tor匿名通讯系统虽然使用了多种混淆加密技术，但是其数据的数据包到达间隔时间、流持续时间和流变为空闲态之前活动时间等时间特征与正常流量存在较大差异。本文在UNB-CIC Tor 数据进行Tor网流量特征提取与流量检测实验，并使用改进的深度森林进行训练与学习。当流超时值为16s取得最佳实验结果，在测试集Tor检测的准确率高达99.86%，改进的深度森林较原始深度深林用更少的计算资源取得了更好的预测性能，即改进的深度森林在取得和原始深度森林科比甚至更好取得预测准确率。本文提出的Tor匿名流量检测系统虽然可以准确的检测Tor流量，但是没有实现在线检测功能，未来将在改方面进行研究与改进。

参考文献

[1] Dingledine R, Mathewson N, Syverson P. Tor: The second-generation onion router[R]. Naval Research Lab Washington DC, 2004, 46(4): 337-337.

[2] Saputra F A, Nadhori I U, Barry B F. Detecting and blocking onion router traffic using deep packet inspection[C]//2016 International Electronics Symposium (IES). IEEE, 2016: 283-288.

[3] Lashkari A H, Draper-Gil G, Mamun M S I, et al. Characterization of tor traffic using time based features[C]//ICISSp. 2017: 253-262.