面向分布式拒绝服务攻击检测的分类算法评估

摘要：分布式拒绝服务攻击旨在利用恶意流量耗尽目标网络的资源，这会对服务的可用性造成威胁。随着互联网的发展，在过去二十年里，研究人员提出了许多检测系统，特别是入侵检测系统（IDS），尽管用户和组织机构不断地发现其应对 DDoS 时所面临的挑战和威胁。虽然，IDS 是保护关键网络免遭不断变化的入侵活动干扰的第一道防线，然而它应该一直保持更新以检测任何异常行为，以便于保护服务的完整性、机密性和可用性。但是，新的检测方法、技术、算法的准确率很大程度上依赖于精心设计的数据集的存在。这些数据集用于训练和利用构建的分类模型进行评估。在本文中，我们使用主要的监督分类算法进行实验，以准确地将 DDoS 和合法流量进行分类。在所有的分类器中，基于树的分类器和基于距离的分类器所表现出的性能最优。

关键词：机器学习，分布式拒绝服务，逻辑回归，朴素贝叶斯，支持向量机，决策树，随机森林，K-近邻算法。

I. 引言

DDoS 攻击已经成为最严重的网络入侵行为之一，对计算机网络的基础设施和各种基于网络的服务产生很严重的威胁[1]。它们非常显著，由于这些攻击很容易发起，会对组织机构造成巨大损失，而且很难追溯并找到真正的攻击者。DDoS 攻击通过耗尽网络资源来破坏网络的可用性,从而导致服务被拒绝。在过去的几年里,这类攻击的数量和规模在迅速增加。随着数据量的增大，攻击持续时间更短的趋势变得越来越普遍[6]。大多数现有的工作都使用像 KDD Cup’99 数据集[2]或 DARPA 数据集[3]来检测 DDoS 攻击。然而，随着时间的流逝，网络犯罪和攻击在以一种巧妙的方式入侵目标网络环境。因此，使用包含各种新型攻击特征的最新数据集来训练分类器将会改善分类器的性能。我们正在利用 CICDDoS2019 数据集来完成我们的分析[4]。

我们工作的目标是实现多个监督分类器，利用 CICDDoS2019 数据集来训练模型以检测 DDoS 攻击。我们关注的焦点是以更高的准确率减少误报，最终改善生产系统的正常运行时间和组织机构的声望。

II. 背景和相关工作

通过从 web 服务器日志中获取特征，如平均数据报大小、传入比特率和传出比特率、源 IP 地址和目的 IP 地址及其端口等[5]，可以用来检测网络流量是否存在异常。拒绝服务攻击主要有两种类型。第一种类型是网络层 DoS 攻击，该类攻击会耗尽网络

资源，从而使实际使用的用户无法连接。第二种攻击类型是应用层 DoS 攻击。该类攻击会耗尽网络资源，使合法用户的请求被拒绝。在 DDoS 攻击中，攻击者会控制多台称为“僵尸机”的机器，攻击者从这些机器上运行被称为“僵尸代码”的脚本，从而攻击目标服务器。

攻击主要分为两大类，一类是反射攻击，一类是利用攻击。在反射攻击中，攻击者的身份保持隐藏，而在利用攻击中，情况并非如此。这两类攻击都能利用应用层协议、传输层协议或者两者的结合来实现。基于 TCP 协议的反射攻击包括 MSSQL、SSDP，而基于 UDP 的反射协议包括 CharGen、NTP、TFTP。

Kurniabudi 在[7]中已经分析了巨大网络流量的相关和重要特征。Ring 等人已经确定了 15 种不同属性来评估各个数据集的适用性[8]。Idhammad 描述了基于网络熵估计、集群、信息增益比和树算法的半监督机器学习方法[9]。研究人员在[10]提出了 INDB（利用朴素贝叶斯进行入侵检测）机制来检测入侵数据报。使用贝叶斯算法的原因是它的可预测性特征。一个被广泛使用的 IDS 分类算法被 Alenezi 和 Reed 在[11]中提出，讨论了 DoS/DDoS 攻击的难点和特征，并且使用三种不同的分类算法来分析数据。Alpna 和 Malhotra 在 KNN 和随机森林算法的帮助下提出了检测 DDoS 攻击的架构[12]。Singh 等人开发了一种改进的 SVM 算法来检测网络攻击[13]。现有的许多相关工作均涉及到 DDoS 攻击检测。然而，大多数研究基于特定的分类算法进行数据集的评估。而且尝试使用一些过时的数据集，如：KDDCup’99 数据集[2]或 DARPA 数据集[3],来实现性能的优化[14-16]。而在本文中,我们使用最新数据集 CICDDoS2019[5]在六个不同的分类算法上进行比较分析。

III. 数据集和方法论

数据集由 7 个 csv 文件组成，包含超过 10GB 的数据。我们利用特征提取算法找到最重要的特征，然后执行数据预处理技术，例如：数据清洗、归一化和无穷大值的消除。一旦模型准备好之后，就可以使用测试集来评估模型的准确率、精确率、召回率、F1 分数、真正例和真负例。如果准确率不够理想，则对每种分类算法进行了优化。除此之外，还分析了训练集和测试集的分割比。

DDoS 攻击通常通过僵尸网络或多个僵尸设备进行。因此，在目标服务器接收数据报时，会有多个 IP 地址或 MAC 地址。但是，通过分析数据报长度、流量持续时间、前向数据报总数等属性，可以帮助我们识别对应请求是合法请求还是恶意请求。为了

比较数据报，可以应用数据挖掘技术来计算概率或发生次数，以分类数据报。我们利用如下六种机器学习算法来对恶意流量进行分类：逻辑回归、支持向量机、朴素贝叶斯、K-近邻、决策树和随机森林算法。

我们使用由新不伦瑞克大学构建的包含 88 个特征的数据集进行实验。该数据集可在加拿大网络安全研究所网站上公开获取[5]。该数据集收集了不同类型的攻击数据，如 Portmap, LDAP, MSSQL, UDP, UDPLag 等。如果请求来自于一个合法用户，那么就会被标记为‘Benign’，否则会被标记为特定攻击名称。该数据集是出于分析的目的而创建的，并按天组织。CIC 每天都会记录原始数据，包括来自每台服务器的网络流量和事件日志。真正的数据集多于 88 个特征，但是 CIC 进行了降维，他们使用 CICFlowmemter-V3[17]，并产生了最重要的 88 个特征进行分析，并提供了 csv 文件。如果有人想要自行提取特征，他们也共享了 PCAP 文件。

我们基于该数据集进行两种类型的实验。首先，我们从每一个 csv 文件中随机采样 30000 行，共计 200000 行作为我们的数据分析样本，这是一个不平衡的数据集。在第二类实验中，我们从每一个 csv 文件中选取数量相等的合法数据和攻击数据元组，从而构建一个完全平衡的训练和测试数据集。

表 1 显示了每个文件中总记录数与正常类别（例如，标签为‘BENIGN’）的对比。更多有关数据集的详细信息参见[18]。数据集中的 IP 地址在训练模型之前被转换为数值整数。

CSV File Name	Total Rows	Benign Rows
LDAP	2113234	5124
MSSQL	5775786	2794
NetBIOS	3455899	1321
Syn	4320541	35790
UDP	3782206	3134
UDPLag	725165	4068
Portmap	191694	4734
Total	**20364525**	**56965**

表 1　标签数据的分布

我们选择了单变量特征选择技术。这是一种统计测试，能够被用来选择那些和输出标签关系最密切的特征。scikit-learn 库中提供了 SelectKBest 类以帮助我们实现该算法，并给出与我们的分类标签最相关的特征的结果。我们使用最相关的 25 种特征来训练我们的模型。我们使用基于树分类器自带的 Feature Importance 类来获取数据集中每种特征的重要性。图 1 展示了 15 种最重要的特征。

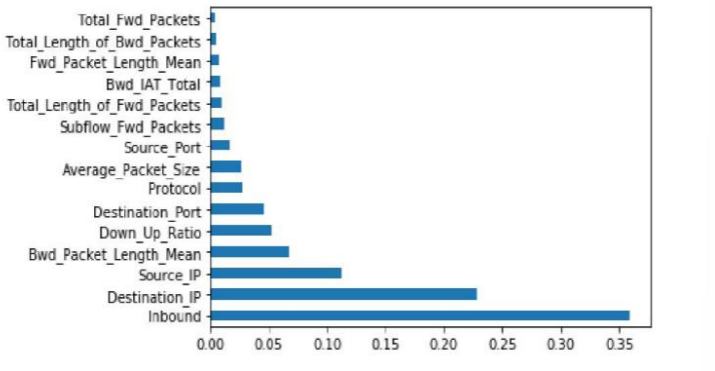

图 1 变量重要性

IV. 实验结果与结论

A. 评估指标

为了评估分类器性能，我们使用基于混淆矩阵的主要性能指标。该矩阵包含由机器学习模型得到的真实和预测分类信息。为了公平起见，我们也在结果表格中包含了 TP、TN、FP 和 FN 的值。正如在之前章节中提到的，我们基于不平衡数据集和平衡数据集分别实现了六种不同的机器学习分类算法。我们使用 python 中的 scikit-learn 库实现了上述两种技术。

B. 实验

我们从每个 csv 文件中随机采样，采样值分别为 30K、40K、50K，来衡量合法流量和攻击流量的比率。真实数据集会包含更少的合法流量数据，而采样本身存在偏差。当使用不平衡数据进行模型训练时，平均来看，与攻击标签相比，会有 0.5%到 0.7%是合法流量。表 2 显示了类别分布。

Sample	Attack (1)	Benign (0)
30K	208710	1263
40K	278302	1698
50K	347780	2220

表 2 类别标签分布（当数据集随机筛选时）

为了避免分类模型准确率上存在的偏差，我们也构建了平衡数据集。我们分别从 7 个 csv 文件中选择全部合法流量元组，然后随机采样相同数量的攻击流量。最终，我们从所有文件中收集了 105042 行数据，其中，合法数据和攻击数据数量相等。由于数据量非常小，我们将相同的数据再次添加到现有数据框中，为了将训练集的大小增加到超过 200K 行，从而与不平衡数据集的数据规模相当。

C. 结果

每个分类器都使用准确率和其他评估指标，例如精确率、召回率和 F1 分数，来进行评估。表 3 显示了在不平衡数据集下每种分类算法的总体准确率；表 4 显示了在平衡数据集下的输出结果。表中数据为五轮实验中的最优值。

						macro avg		
Unbalanced Dataset	**TP**	**TN**	**FN**	**FP**	**Accuracy**	**Precession**	**Recall**	**F1 Score**
Decision Tree	62599	398	3	0	99.99523	1	1	1
Naive Bayes	61199	370	31	1400	97.72857	0.6	0.95	0.66
Logistic Regression	62619	213	164	4	99.73333	0.99	0.78	0.86
Support Vector Machine	62663	0	337	0	99.46507	0.5	0.5	0.5
K Nearest Neighbor	62598	401	0	1	99.99841	1	1	1
Random Forest	62602	397	0	1	99.99841	1	1	1

表 3 不平衡数据集实验结果

						macro avg		
Balanced Dataset	**TP**	**TN**	**FN**	**FP**	**Accuracy**	**Precession**	**Recall**	**F1 Score**
Decision Tree	31577	31449	0	0	100	1	1	1
Naive Bayes	31387	29278	2290	71	96.25392	0.96	0.96	0.96
Logistic Regression	31276	8730	12819	201	79.34185	0.85	0.79	0.78
Support Vector Machine	31577	0	31449	0	50.10154	0.25	0.5	0.33
K Nearest Neighbor	31477	31549	0	0	100	1	1	1
Random Forest	31477	31549	0	0	100	1	1	1

表 4 平衡数据集实验结果

由于不平衡数据集偏向于攻击数据，因此所有分类算法的准确率是非常高的。但是，这无助于针对 DDoS 攻击检测选择性能最优的分类算法。在数据不平衡的情况下，除了朴素贝叶斯算法外，其他所有算法的性能都表现地很好。相反，平衡数据集的准确率变化很小。如表 4 所示，基于树的算法，如决策树、随机森林和基于距离的分类算法 K-NN 表现最优，朴素贝叶斯取得较好的性能，而剩余的分类算法 SVM 和逻辑回归则表现较差。图 2 显示了在不平衡数据集和平衡数据集下每一种分类算法的准确率。除此之外，图 3、图 4 和图 5 分别显示了两种数据集下精确率、召回率和 F1 分数的比较。

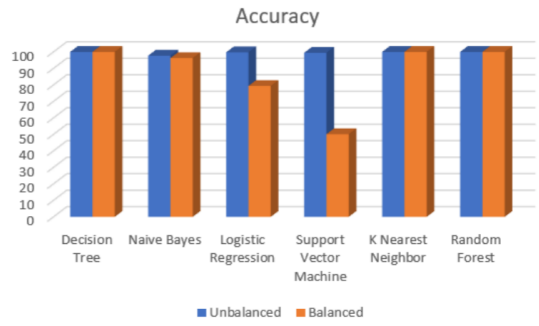

图 2 分类算法准确率

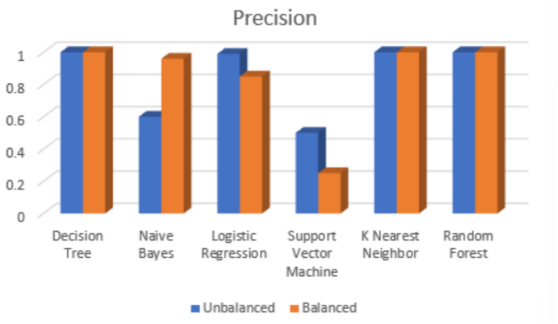

图 3 分类算法精确率

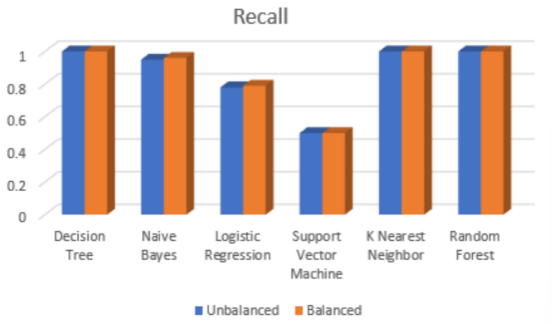

图 4 分类算法召回率

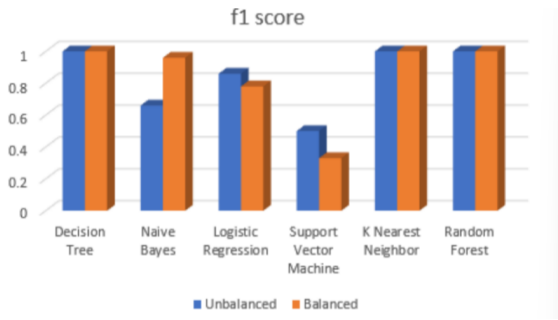

图 5 分类算法 F1 分数

在分析输出结果后，基于树的分类算法，如决策树、随机森林和基于距离的分类算法，在两种数据集上均性能最优，准确率几乎达到 100%。甚至将其他指标考虑进去，这三类分类算法仍然性能最优。然而，当每个分类器的参数发生变化时，可以注意到性能的微小变化。我们试图为每一种算法提供最优性能。

V. 未来工作建议

虽然初步实验结果鼓舞人心，但是该项工作还可以在如下方面进行扩展：a）由于硬件设备的局限性，我们仅仅使用略多于 200000 行的数据进行实验。后续，我们会计划选择多于 1 百万数据的数据集进行实验。这样将为我们提供更准确的预测模型。b）我们会基于每一种不同的 DDoS 攻击来进行数据挖掘。因为很有可能 Portmap 能够被 K-NN 很好地检测出来，而 UDPlag 可能利用朴素贝叶斯效果更好。如果这一点能够得到证实，我们会将所有模

型融合为一个模型，使得针对所有种类的 DDoS 攻击准确率接近 100%。 c）我们会尝试不同的特征选择技术。

VI. 结论

在本文中，我们使用 CICDDoS2019 数据集，这是一个最新的数据集，其中包含最新的 DDoS 攻击特征。我们使用主要的监督分类算法进行实验，以正确地区分攻击和合法流量。和其他算法相比，决策树、随机森林和 K-NN 性能最优。虽然初步实验结果鼓舞人心，但是我们计划在扩展后的数据集上以及针对不同类型的 DDoS 攻击进行扩展工作。在未来工作中，我们将重点关注这些方向。

参考文献

[1]Neustar.(2014).2014-Neustar Annual DDoS Attacks and Impact Report[Online].Available:http://neustar.biz/resources/whitepapers/ddos-protection/2014-annual-ddos-attacks-andimpact-report.pdf

[2]KDDCUP’99-http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[3]DARPA Dataset [Online].Available:https://www.ll.mit.edu/r-d/datasets

[4]I.Sharafaldin, A.Habibi L.Saqib Hakak, and A. Ghorbani, "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy", IEEE 53rd International Carnahan Conference on Security Technology, Chennai, India, 2019.

[5] CICDDoS2019 - https://www.unb.ca/cic/datasets/ddos-2019.html Last accessed August 2020.

[6] L Buczak, and E.Guven. "A survey of data mining and machine learning methods for cyber security intrusion detection." IEEE Communications surveys & tutorials 18, no. 2 (2015): 1153-1176.

[7] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Idris, A. M. Bamhdi and R. Budiarto, "CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection," in IEEE Access, vol. 8, pp.132911-132921, 2020, doi: 10.1109/ACCESS.2020.3009843.

[8] M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho. "A survey of network-based intrusion detection data sets." Computers & Security 86 (2019):147-167.

[9] M. Idhammad, K. Afdel, and M. Belouch. "Semi-supervised machine learning approach for DDoS detection." Applied Intelligence 48, no. 10 (2018): 3193-3208.

[10] V. Hema and C. E. Shin. "DoS Attack Detection Based on Naive Bayes Classifier." Middle-East Journal of Scientific Research 23 (2015): 398-405.

[11] M. Alenezi and M. J. Reed, “Methodologies for detecting dos/ddos attacks against network servers,”in 7th Intl Conference on Systems and Networks Communications, ICSNC SemiMarkov models, 2012.

[12] Alpna and S. Malhotra,“DDoS Attack Detection and Prevention Using Ensemble Classifier (RF)”, IJARCSSE, 2016.

[13] S.Singh,S.Agrawal, M. A. Rizvi, and R.S. Thakur. "Improved Support Vector Machine for Cyber Attack Detection." In Proceedings of the World Congress on Engineering and Computer Science, vol. 1. 2011.

[14] Wankhede, S., & Kshirsagar, D. (2018). DoS Attack Detection Using Machine Learning and Neural Network. 2018 Fourth International Conference on Computing Communication Control and Automation.

[15] Shuyuan Jin and D. S. Yeung, "A covariance analysis model for DDoS attack detection," 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577), Paris, France, 2004, pp. 1882-1886 Vol.4, doi: 10.1109/ICC.2004.1312847.

[16] P. Khuphiran, et al.,. (2018). Performance Comparison of Machine Learning Models for DDoS Attacks Detection. 2018 22nd International Computer Science and Engineering Conference

[17]CICFlowMeter-v3 codebase can be accessed using this link - https://github.com/ahlashkari/CICFlowMeter

[18] Iman Sharafaldin, Arash Habibi Lashkari, Saqib Hakak, and Ali A. Ghorbani, "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy", IEEE 53rd International Carnahan Conference on Security Technology, Chennai, India, 2019.

[19] S. Kumar and B. Xu. "A Machine Learning Based Approach to Detect Malicious Fast Flux Networks." In 2018 IEEE Symposium Series on Comp. Intelligence, pp. 1676-1683.

[20] J. Chelladhurai, P. Chelliah, and S. Kumar. "Securing docker containers from denial of service (dos) attacks."2016 IEEE Intl Conference on Services Computing (SCC), pp. 856-859.

[21] C. Chrane and S. Kumar. "An Examination of Tor Technology Based Anonymous Internet." In I n SITE 2015: Informing Science+ IT Education Conferences: USA, pp. 145-153. 2015.

[22] S. Srinivasan, S,and S. P. Alampalayam."Intrusion Detection Algorithm for MANET." International Journal of Information Security and Privacy (IJISP) 5, no. 3 (2011): 36-49.

[23] S. Kumar, "Classification and review of security schemes in mobile computing." Wireless Sensor Network 2, no. 06 (2010): 41

[24] S. Alampalayam, and A. Kumar. "Predictive security model using data mining." In IEEE GLOBECOM'04, vol. 4, pp. 2208-2212.

[25] S. Alampalayam, and A. Kumar. "An adaptive and predictive security model for mobile ad hoc networks." Wireless Personal Communications 29, no. 3-4 (2004): 263-281.

实验结果

Balanced Dataset	Accuracy	Precession	Recall	F1 Score
Decision Tree	0.995	0.991	1	0.995
Naïve Bayes	0.995	0.991	0.998	0.995
Logistic Regression	0.879	0.856	0.913	0.883
K Nearest Neighbor	0.999	0.999	0.999	0.999
Random Forest	0.998	0.995	0.999	0.998

实验所用字段及含义

本次实验按照论文逻辑选择最相关的 25 种特征构建数据集。以下是字段含义。

'Unnamed: 0'：数据集第一个字段
'Flow ID'：数据流标识符，区分不同的数据流
' Source IP'：源 IP 地址，即发出数据流的设备 IP 地址
' Source Port'：源设备使用的端口号
' Destination IP'：目的 IP 地址，即接收数据流的设备 IP 地址
' Destination Port'：目的设备使用的端口号
' Timestamp'：时间戳，表示数据流的时间
' Fwd Packet Length Min'：表示源设备所发送的最小数据包长度
'Bwd Packet Length Max'：表示目的设备所发送的最大数据包长度
' Bwd Packet Length Min'：表示目的设备所发送的最小数据包长度
' Bwd Packet Length Mean'：表示目的设备所发送的平均数据包长度
'Flow Bytes/s'：表示数据流每秒传送的字节数
' Flow Packets/s'：表示数据流每秒传送的数据包数
'Fwd PSH Flags'：表示前向 PSH 标志

'Fwd Packets/s'：表示每秒源设备发送的数据包数
' Min Packet Length'：数据流中的最小数据包长度
' Packet Length Std'：数据包长度的标准差
' RST Flag Count'：数据流中的 RST 标志数量
' ACK Flag Count'：数据流中的 ACK 标志数量
' URG Flag Count'：数据流中的 URG 标志数量
' CWE Flag Count'：数据流中的 CWE 标志数量
' Down/Up Ratio'：数据流中下行流量与上行流量的比率
' Avg Fwd Segment Size'：表示源设备发送的平均段大小
' Avg Bwd Segment Size'：表示目的设备发送的平均段大小
' Inbound'：表示是否为入站流量