(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 115051874 B
(45)授权公告日 2022.12.09

(21)申请号 202210916554.0
(22)申请日 2022.08.01
(65)同一申请的已公布的文献号
申请公布号 CN 115051874 A
(43)申请公布日 2022.09.13
(73)专利权人 杭州默安科技有限公司
地址 311100 浙江省杭州市余杭区仓前街道余杭塘路2616号3号楼1楼
(72)发明人 王乐
(74)专利代理机构 杭州裕阳联合专利代理有限公司 33289
专利代理师 何宇梁
(51)Int.Cl.
*H04L 9/40*(2022.01)
*G06K 9/62*(2022.01)

(56)对比文件
CN 114172748 A,2022.03.11
US 2020236131 A1,2020.07.23

审查员 张燕燕

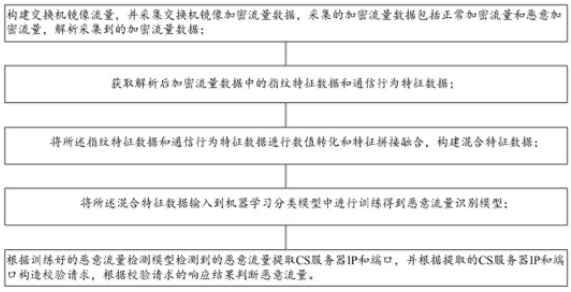
权利要求书2页 说明书6页 附图1页

(54)发明名称

一种多特征的CS恶意加密流量检测方法和系统

(57)摘要

本发明公开了一种多特征的CS恶意加密流量检测方法和系统，所述方法包括：构建交换机镜像流量，并采集交换机镜像加密流量数据，采集的加密流量数据包括正常加密流量和恶意加密流量，解析采集到的加密流量数据；获取解析后加密流量数据中的指纹特征数据和通信行为特征数据；将所述指纹特征数据和通信行为特征数据进行数值转化和特征融合，构建混合特征数据；将所述混合特征数据输入到机器学习分类模型中进行训练得到恶意加密流量检测模型；根据训练好的恶意加密流量检测模型检测到的恶意流量提取CS服务器IP和端口，并根据提取的CS服务器IP和端口构造校验请求，根据校验请求的响应结果判断恶意流量。




CN 115051874 B

1.一种多特征的CS恶意加密流量检测方法,其特征在于,所述方法包括:

构建交换机镜像流量,并采集交换机镜像加密流量数据,采集的加密流量数据包括正常加密流量和恶意加密流量,解析采集到的加密流量数据;

获取解析后加密流量数据中的指纹特征数据和通信行为特征数据;

将所述指纹特征数据和通信行为特征数据进行数值转化和特征拼接融合,构建混合特征数据;

将所述混合特征数据输入到机器学习分类模型中进行训练得到恶意加密流量检测模型;

根据训练好的恶意加密流量检测模型检测到的恶意流量提取CS服务器IP和端口,并根据提取的CS服务器IP和端口构造校验请求,根据校验请求的响应结果判断恶意流量;

获取所述恶意加密流量数据和正常加密流量数据后,提取对应的字符串类型特征数据,并将所述字符串类型数据根据字符串的类型和字符串长度进行赋值,根据字母的长度进行字符串的赋值,赋值范围设置为0-100,对于非字母的字符串类型根据非字母的字符串类型在101-200范围内赋值,非字母字符串类型赋值时通过加100的基础赋值的方式进行不同字符串类型的分段赋值;

数据拼接是基于单一维度的拼接方式,若指纹特征F = {f1,f2,……,fn},通信行为特征 T = {t1,t2,……,tn},则拼接之后的特征Feature = F+T = {f1,f2,……,fn,t1,t2,……,tn},其中+为拼接操作,Feature为拼接之后的数值类型特征,fn为指纹特征中的任意一个特征元素,tn为通讯行为特征中任意一个特征元素。

2.根据权利要求1所述的一种多特征的CS恶意加密流量检测方法,其特征在于,所述解析采集到的加密流量数据方法包括:获取加密流量数据后,通过特征匹配的方式将所述流量数据分类为有效流量和无效流量。

3.根据权利要求2所述的一种多特征的CS恶意加密流量检测方法,其特征在于,在完成所述加密流量数据分类后,对有效的正常加密流量数据和恶意加密流量数据分别进行指纹特征提取和通信行为特征提取,其中所述指纹特征包括CS客户端指纹特征和CS服务器指纹特征。

4.根据权利要求1所述的一种多特征的CS恶意加密流量检测方法,其特征在于,所述通信行为特征信息包括五元组信息,其中所述五元组信息分别为源ip、目的ip、源端口、目的端口和通信协议类型。

5.根据权利要求1所述的一种多特征的CS恶意加密流量检测方法,其特征在于,在获取到指纹特征数据和通信行为特征数据后,将所述指纹特征数据和通信行为特征数据中的字符串类型数据根据字符串类型和长度转化为数值型数据。

6.根据权利要求1所述的一种多特征的CS恶意加密流量检测方法,其特征在于,所述混合特征数据的构建方法包括:获取所述指纹特征数据和通信行为特征数据中数值型数据以及转化后的数值型数据,将指纹特征数据所有的数值型数据和所述通信行为中的数值型数据进行拼接,得到拼接后的数值型特征数据。

7.根据权利要求6所述的一种多特征的CS恶意加密流量检测方法,其特征在于,获取拼接后的数值型特征数据,将所述数值型特征数据通过如下公式进行归一化处理:

$$Y=\frac{x-min}{max-min}$$

其中Y为归一化处理的结果，x为拼接后数值型特征数据的任意一个元素，max为拼接后数值型特征数据的最大值，min为拼接后数值型特征数据的最小值。

8.根据权利要求1所述的一种多特征的CS恶意加密流量检测方法，其特征在于，所述恶意加密流量检测模型识别恶意加密流量后，根据所述恶意加密流量的五元组获取CS服务器IP和端口，并生成符合checksum8校验的字符串，将校验的字符串作为访问地址构建校验请求，若满足校验，构造特殊请求下载cs服务器的加密配置文件，并返回校验状态码。

9.一种多特征的CS恶意加密流量检测系统，其特征在于，所述系统执行权利要求1-8中任意一项所述的一种多特征的CS恶意加密流量检测方法。

10.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，所述计算机程序可被处理器执行权利要求1-8中任意一项所述的一种多特征的CS恶意加密流量检测方法。

# 一种多特征的CS恶意加密流量检测方法和系统

## 技术领域

[0001] 本发明涉及网络安全检测技术领域，特别涉及一种多特征的CS恶意加密流量检测方法和系统。

## 背景技术

[0002] 随着信息化进程的不断深入，给人们的社会生活带来巨大便利的同时也引发了严峻的安全挑战。一方面，各种类型的网络攻击层出不穷，大量数据泄露事件频繁出现；另一方面，流量混淆工具的广泛使用使得一些网络流量审查机制疲于应对。随着SSL和TLS协议的广泛使用，人们在受益于加密流量提供的安全通信之余也给恶意软件隐藏他们的恶意行为提供了便利。根据相关报告显示，利用加密通信绕过传统安全控制的网络攻击数量迅速增加，同时近四分之一的恶意软件现在使用TLS进行通信。加密流量的普遍性以及有效载荷的不可见性给网络安全带来了巨大的挑战。因此，现阶段对识别隐藏在加密流量中的恶意活动具有迫切需求和重要意义。目前主流的攻击检测手段有两种，一种是解密后检测，另外一种是不解密检测。业界网关设备主要使用解密流量的方法检测攻击行为，但这种解密方法会消耗大量的资源，成本很高，同时也违反了加密的初衷，解密过程会受到隐私保护相关法律法规的严格限制，出于保护用户隐私的考量，解密后检测因为可能存在泄密风险而会谨慎使用。

## 发明内容

[0003] 本发明其中一个发明目的在于提供一种多特征的CS恶意加密流量检测方法和系统，方法和系统可以对CS工具的恶意加密流量进行检测。通过对CS客户端加密流量的检测可以有效阻止黑客利用CS工具进行渗透测试。

[0004] 本发明另一个发明目的在于提供一种多特征的CS恶意加密流量检测方法和系统，所述方法和系统利用流量解析工具得到流量数据包中包含客户端加密套件和扩展信息等指纹信息，以及CS客户端和服务端通讯相关的行为信息作为判断恶意加密流量的判断依据，将上述指纹信息和行为信息拼接为融合的多特征进行恶意加密流量判断，从而可以通过多维度提高判断的准确率。

[0005] 本发明另一个发明目的在于提供一种多特征的CS恶意加密流量检测方法和系统，所述方法和系统在获取拼接后融合的混合特征输入到机器学习分类算法中进行模型训练，得到用于精确检测恶意加密流量的检测模型，根据训练好的检测模型得到CS行为数据查找到对应的五元组信息，并根据所述五元组信息获取CS服务器IP和端口，通过所述CS服务器IP和端口构造特定访问地址进行校验，并输出校验后的CS服务器状态码作为鉴别恶意加密流量的依据。

[0006] 为了实现至少一个上述发明目的，本发明进一步提供一种多特征的CS恶意加密流量检测方法，所述方法包括：

[0007] 构建交换机镜像流量，并采集交换机镜像加密流量数据，采集的加密流量数据包

括正常加密流量和恶意加密流量，解析采集到的加密流量数据；

[0008] 获取解析后加密流量数据中的指纹特征数据和通信行为特征数据；

[0009] 将所述指纹特征数据和通信行为特征数据进行数值转化和特征融合，构建混合特征数据；

[0010] 将所述混合特征数据输入到机器学习分类模型中进行训练得到恶意加密流量检测模型；

[0011] 根据训练好的恶意加密流量检测模型检测到的恶意加密流量提取CS服务器IP和端口，并根据提取的CS服务器IP和端口构造校验请求，根据校验请求的响应结果判断恶意加密流量。

[0012] 根据本发明其中一个较佳实施例，所述解析采集到的加密流量数据方法包括：获取加密流量数据后，通过特征匹配的方式将所述流量数据分类为有效流量和无效流量。

[0013] 根据本发明另一个较佳实施例，在完成所述加密流量数据分类后，对有效的所述正常加密流量数据和恶意加密流量数据分别进行指纹特征提取和通信行为特征提取，其中所述指纹特征包括CS客户端指纹特征和CS服务器指纹特征。

[0014] 根据本发明另一个较佳实施例，所述通信行为特征信息包括五元组信息，其中所述五元组信息分别为源ip、目的ip、源端口、目的端口和通信协议类型。

[0015] 根据本发明另一个较佳实施例，在获取到指纹特征数据和通信行为特征数据后，将所述指纹特征数据和通信行为特征数据中的字符串类型数据根据字符串类型和长度转化为数值型数据。

[0016] 根据本发明另一个较佳实施例，所述融合特征数据的构建方法包括：获取所述指纹特征数据和通信行为特征数据中数值型数据以及转化后的数值型数据，将指纹特征数据所有的数值型数据和所述通信行为中的数值型数据进行拼接，得到拼接后的数值型特征数据。

[0017] 根据本发明另一个较佳实施例，获取拼接后的数值型特征数据，将所述数值型特征数据通过如下公式进行归一化处理：

[0018] $$Y=\frac{x-\min}{\max-min}$$

[0019] 其中Y为归一化处理的结果，x为拼接后数值型特征数据的任意一个元素，max为拼接后数值型特征数据的最大值，min为拼接后数值型特征数据的最小值。

[0020] 根据本发明另一个较佳实施例，所述恶意加密流量检测模型识别恶意加密流量后，根据所述恶意加密流量的五元组获取CS服务器IP和端口，并生成符合checksum8校验的字符串，将校验的字符串作为访问地址构建校验请求，若满足校验，构造特殊请求下载CS服务器的加密配置文件并返回校验状态码。

[0021] 为了实现至少一个上述发明目的，本发明进一步提供一种多特征的CS恶意加密流量检测系统，所述系统执行上述一种多特征的CS恶意加密流量检测方法。

[0022] 本发明进一步提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序可被处理器执行上述一种多特征的CS恶意加密流量检测方法。

## 附图说明

[0023] 图1显示的是本发明一种多特征的CS恶意加密流量检测方法流程示意图。

[0024] 图2显示的是本发明一种多特征的CS恶意加密流量检测系统结构示意图。

[0025] 图3显示的是本发明中混合特征构建和加密流量检测示意图。

## 具体实施方式

[0026] 以下描述用于揭露本发明以使本领域技术人员能够实现本发明。以下描述中的优选实施例只作为举例，本领域技术人员可以想到其他显而易见的变型。在以下描述中界定的本发明的基本原理可以应用于其他实施方案、变形方案、改进方案、等同方案以及没有背离本发明的精神和范围的其他技术方案。

[0027] 可以理解的是，术语“一”应理解为“至少一”或“一个或多个”，即在一个实施例中，一个元件的数量可以为一个，而在另外的实施例中，该元件的数量可以为多个，术语“一”不能理解为对数量的限制。

[0028] 请结合图1-图3，本发明公开了一种多特征的CS恶意加密流量检测方法和系统，所述方法包括：首先需要对交换机的流量数据进行镜像复制，利用流量捕捉分析工具对所述镜像复制的流量进行捕捉，其中所述流量捕捉分析工具可以采用包括但不仅限于现有的joy流量捕捉分析工具，进一步通过所述流量捕捉分析工具采集待检测的TLS数据流量，其中所述待检测的TLS数据流量包括正常加密流量和恶意加密流量，其中本发明通过特征匹配的方式筛选出所述待检测流量中有效的TLS加密数据流量。其中所述特征匹配的方式主要是检查CS(加密套件)字段是否存在以及所含的加密套件支持的算法数量是否大于0，如果不符合该条件，则认为该加密数据不包含握手信息，不包含后期检测需要的指纹特征，此时该加密流量为无效流量。正常数据的采集主要是通过在内网搭建环境抓取内网https流量，此时采集的流量数据认为是正常https流量。CS恶意加密流量的采集也是搭建不同版本的CS客户端服务器构造客户端与服务器交互环境采集的数据，此时采集的流量数据认为是CS恶意流量。

[0029] 进一步的，本发明在完成所述样本数据的构建后，进一步对所述样本数据进行特征提取，用于构建多特征的混合特征数据，其中本发明需要分别对所述样本数据中的正常加密流量数据和恶意加密流量数据进行特征提取，分别提取所述正常加密流量数据和恶意加密流量数据的指纹特征数据和通信行为特征数据，其中所述指纹特征数据包括但不仅限于CS客户端指纹特征数据和CS服务端指纹特征数据，其中所述指纹数据的类型包括TLS版本、客户端扩展数量、服务端扩展数量、客户端密码套件数量等，所述指纹特征数据用于描述CS客户端和CS服务端对象特征。其中所述通讯行为特征数据包括但不仅限于五元组信息、会话时长，上行包数、下行包数、上行包传输字节数、下行包传输字节数和流持续时长等，其中所述五元组信息包括源ip、目的ip、源端口、目的端口和通信协议类型。本发明中所述通信行为特征数据优选为TLS通信行为特征数据，用于描述TLS通讯相关的特征数据。

[0030] 值得一提的是，在对所述恶意加密流量数据和正常加密流量数据进行指纹特征数据和通信行为特征数据提取后，进一步将所述指纹特征数据和通信行为特征数据进行拼接，并将拼接后的数据进行归一化处理得到混合特征数据，并进一步将所述混合特征数据输入到机器学习分类模型中训练，得到训练好恶意加密流量检测模型。

[0031] 需要说明的是，由于所述恶意加密流量数据和正常加密流量数据均包括数值型特征数据和字符串型特征数据，因此本发明为了便于所述机器学习分类模型的训练，本发明进所述恶意加密流量数据和正常加密流数据中的字符串类型特征数据进行数值转化，转化为数值型特征数据。在本发明其中一个较佳实施例中，获取所述恶意加密流量数据和正常加密流量数据后，提取对应的字符串类型特征数据，并将所述字符串类型数据根据字符串的类型和字符串长度进行赋值，比如字符串类型为字母abc时，可以根据所述字母的长度进行字符串的赋值，赋值范围可以设置为0-100等，对于非字母的字符串类型可以根据非字母的字符串类型在101-200范围内赋值，也就是说非字母字符串类型赋值时可以通过加100的基础赋值的方式进行不同字符串类型的分段赋值，因此本发明通过上述不同类型字符串的分段赋值的方式，可以有效地将所述正常加密流量和恶意加密流量中对应的指纹特征数据和通信行为特征数据的字符串类型数据转化为数值型字符串；因此在进行数值转化后，所述指纹特征数据和通信行为特征数据均为数值型特征数据。因此可以对数值类型的指纹特征数据和通信行为特征数据进行数据拼接。需要说明的是本发明采用的数据拼接是基于单一维度的拼接方式。即数值类型指纹特征数据和数值类型的通信行为特征数据，举例来说：若指纹特征F = {f1,f2,……,fn}，通信行为特征 T = {t1,t2,……,tn}，则拼接之后的特征Feature = F+T = {f1,f2,……,fn,t1,t2,……,tn}，其中+为拼接操作，Feature为拼接之后的数值类型特征。fn为指纹特征中的任意一个特征元素，所述特征元素包括但不仅限于TLS版本、客户端扩展数量、服务端扩展数量、客户端密码套件数量。所述tn为通讯行为特征中任意一个特征元素，所述特征元素包括但不仅限于五元组信息任意一个(源ip、目的ip、源端口、目的端口、通信协议类型)、会话时长，上行包数、下行包数、上行包传输字节数、下行包传输字节数、流持续时长。

[0032] 在本发明另一个较佳实施例中，可以利用parseInt()函数将所述指纹特征和通信行为特征中的字符串转为数值类型，其中parseInt()函数可以将字符串转换为整数数值类型。或者可以利用parseFloat()函数将所述指纹特征和通信行为特征中的字符串转为浮点数值类型，需要说明的是，上述parseInt()函数和parseFloat()函数均要在String类型调用下才能实现字符串到数值类型的转换。

[0033] 在完成转化为数值型的指纹特征数据和通信行为特征数据的拼接后，将拼接后的数值型特征数据进行归一化处理，其中所述归一化处理的方法包括：计算拼接后数值类型特征数据的最大特征值，并计算拼接后数值类型特征数据的最小特征值，利用如下公式计算归一化结果：

[0034] $$Y = \frac{x - min}{max - min}$$

[0035] 其中其中Y为归一化处理的结果，x为拼接后数值型特征数据的任意一个元素，max为拼接后数值型特征数据的最大值，min为拼接后数值型特征数据的最小值。所述归一化处理后的拼接数据作为最终的混合特征数据，用于所述机器学习分类模型的输入数据。本发明中优选的机器学习分类模型为有监督学习算法的XGboost模型。所述XGboost模型训练方法包括：构建XGboost模型分类器，每个分类器为一个节点，并设置XGboost模型的深度；构建XGboost模型的目标函数，本发明中目标函数将根据训练次数设置，根据所述XGboost模型分类器构建决策树，将所述混合特征数据输入到所述决策树中进行残差计算，生成新的

子分类器作为决策树的子叶拟合所述上一个分类器的残差，直到目标函数最小，并利用损失函数计算模型的收敛情况，通过调节所述XGboost模型的包括但不仅限于学习速率(learning_rate)、gamma参数、最大深度(max_depth)、min_child_weight参数等，根据损失函数获取最优的恶意加密流量检测模型。

[0036] 具体而言，所述XGboost模型训练方法详细步骤包括：构建训练数据，训练数据包含设置有正常流量和CS恶意流量的标签，针对每一训练数据的流量，提取其指纹特征和TCP通信行为特征，经过数据预处理之后将两部分特征进行拼接得到最终的用于训练模型的待检测特征数据。对待检测特征数据使用十折交叉验证法将特征数据划分为训练集、测试集和验证集。将划分好的训练集输入到XGboost模型，对Xgboost模型的参数如迭代次数n_estimators、学习率、最大深度等使用网格搜索方法GridSearchCV对模型进行调参，并根据评分指标输出模型最优参数。根据确定好的参数训练模型，对验证集数据进行预测，预测结果与验证集标签对比输出准确率、召回率、F1_score值用于判断模型的效果。本发明利用XGboost模型通过上述训练得到的模型为所述恶意加密流量检测模型，对此其他现有机器学习分类模型训练后作为本发明的恶意加密流量检测模型，对此本发明不再详细赘述。

[0037] 在完成上述基于XGboost模型的恶意加密流量检测模型的训练后，利用所述恶意加密流量检测模型识别恶意流量，并进一步解析所述恶意加密流量，其中解析获取所述恶意加密流量中CS服务器IP和端口，并进一步构造校验请求，其中所述校验请求可以被构造为:https://ip:port/uri，其中该校验请求网址中的IP为检测到的恶意加密流量CS服务器网址，该校验请求网址中的uri为不同操作系统自身生成的符合checksum8校验的字符串，其中通过所述校验请求访问特定的uri后，通过响应信息判断是否满足校验，若返回信息的校验状态码为200，则表示校验请求成功，此时可以拉取CS服务器和CS客户端通信的配置文件，通过该配置文件可以构建伪造的CS客户端和服务器通信。若返回的状态码为404，则代表请求校验的资源不存在等，此时校验失败，其中所述配置文件为CS服务器用于后续渗透操作的配置文件。其中所述校验方法为:计算字符序列的ASCII编码值，得到计算值，设定固定值，判断计算值与固定值是否相等来进行校验，若相等，则认为校验成功，若不相等则校验失败。

[0038] 本发明利用上述方法和系统，测试了攻击者经常使用的常见的不同CS版本以及不同类型的木马(普通木马，无状态木马，powershell木马)均能检出。并且检测的报错率低至0.03%，因此本发明具有检测时间开销小，检测准确率高、误报与漏保率低的优点。请参考图3，本发明中利用模型检测出的恶意流量CS服务器IP和端口，可以通过CS干扰装置阻止CS服务器和客户端对互联网交换机的攻击。其中所述CS干扰装置可以通过CS工具实现，本发明对此不再赘述。

[0039] 特别地，根据本发明公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分从网络上被下载和安装，和/或从可拆卸介质被安装。在该计算机程序被中央处理单元(CPU)执行时，执行本申请的方法中限定的上述功能。需要说明的是，本申请上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是但不限于电、

磁、光、电磁、红外线段、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线段的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线段、电线段、光缆、RF等等，或者上述的任意合适的组合。

[0040] 附图中的流程图和框图，图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0041] 本领域的技术人员应理解，上述描述及附图中所示的本发明的实施例只作为举例而并不限制本发明，本发明的目的已经完整并有效地实现，本发明的功能及结构原理已在实施例中展示和说明，在没有背离所述原理下，本发明的实施方式可以有任何变形或修改。

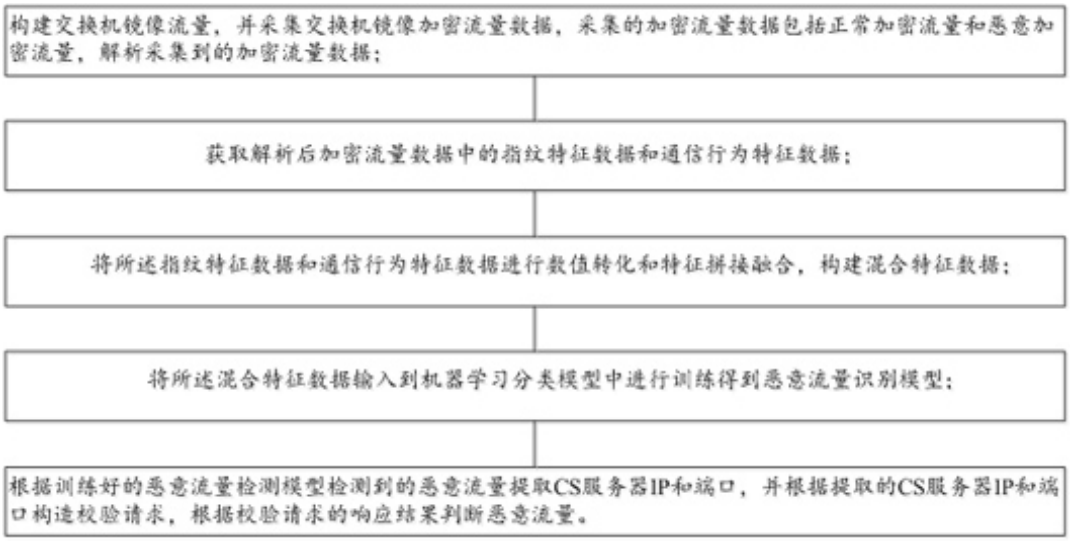

构建交换机镜像流量，并采集交换机镜像加密流量数据，采集的加密流量数据包括正常加密流量和恶意加密流量，解析采集到的加密流量数据；
获取解析后加密流量数据中的指纹特征数据和通信行为特征数据；
将所述指纹特征数据和通信行为特征数据进行数值转化和特征拼接融合，构建混合特征数据；
将所述混合特征数据输入到机器学习分类模型中进行训练得到恶意流量识别模型；
根据训练好的恶意流量检测模型检测到的恶意流量提取CS服务器IP和端口，并根据提取的CS服务器IP和端口构造校验请求，根据校验请求的响应结果判断恶意流量。

图1

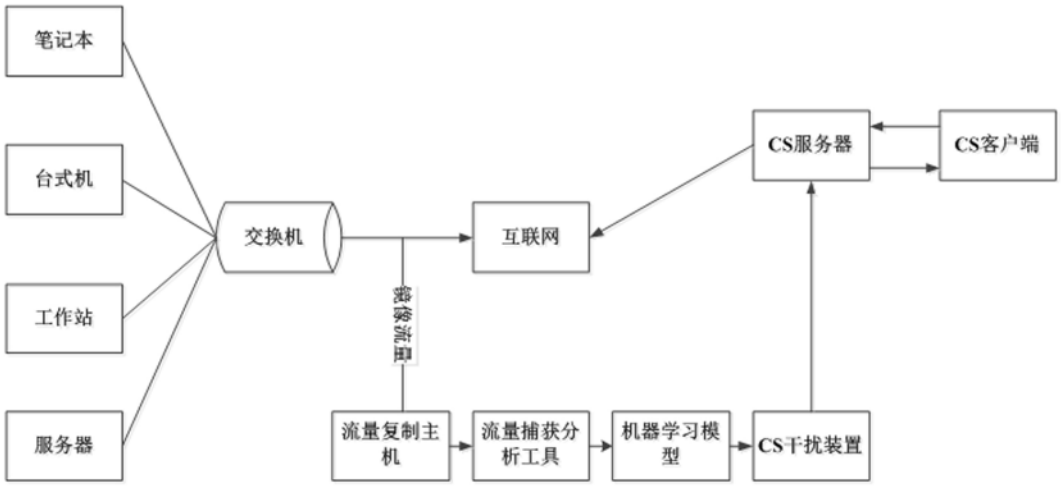

笔记本
台式机
工作站
服务器
交换机
互联网
CS服务器
CS客户端
镜像流量
流量复制主机
流量捕获分析工具
机器学习模型
CS干扰装置

图2

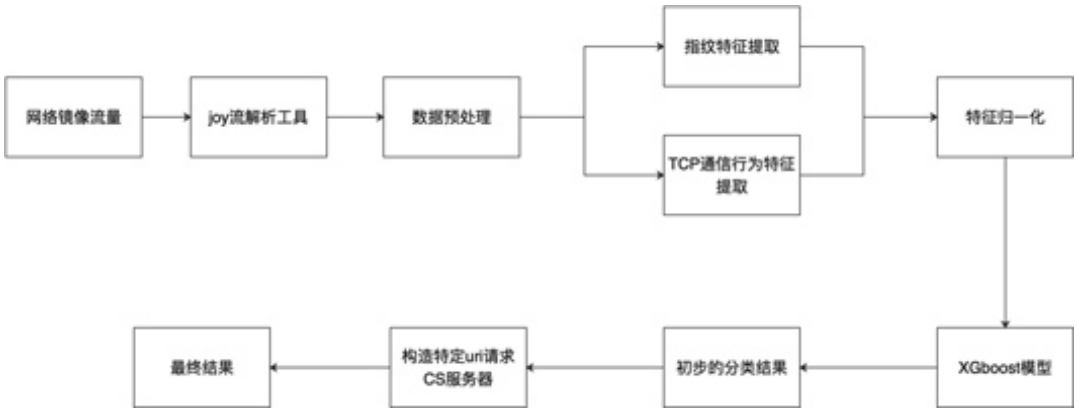

网络镜像流量
joy流解析工具
数据预处理
指纹特征提取
TCP通信行为特征提取
特征归一化
XGboost模型
初步的分类结果
构造特定uri请求CS服务器
最终结果

图3