(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118540077 A
(43)申请公布日 2024.08.23

(21)申请号 202310166669.7

(22)申请日 2023.02.22

(71)申请人 中国科学院计算机网络信息中心
地址 100190 北京市海淀区中关村南四街4号院内2号楼

(72)发明人 龙春 付豪 魏金侠 宫良一 付豫豪 王跃达

(74)专利代理机构 北京知舟专利事务所(普通合伙) 11550
专利代理师 郭韫

(51)Int.Cl.
H04L 9/40(2022.01)
H04L 61/4511(2022.01)

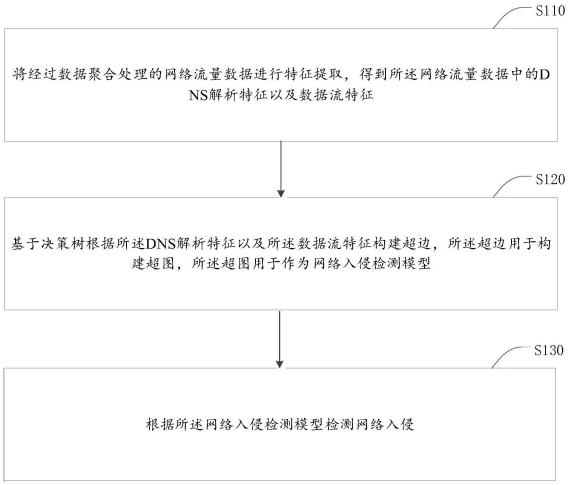
权利要求书2页 说明书8页 附图2页

(54)发明名称

网络入侵检测方法以及装置、存储介质、电子装置

(57)摘要

本申请公开了网络入侵检测方法以及装置、存储介质、电子装置，其中所述方法包括将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征；基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型；根据所述网络入侵检测模型检测网络入侵。通过本申请通过DNS解析行为、流量通信行为综合进行考虑攻击者行为并提取多维特征进行分析，可以更加详细地描述出攻击的整个过程，可以检测出更加隐蔽的攻击。



CN 118540077 A

1.一种网络入侵检测方法,其特征在于,所述方法包括:

将经过数据聚合处理的网络流量数据进行特征提取,得到所述网络流量数据中的DNS解析特征以及数据流特征;

基于决策树根据所述DNS解析特征以及所述数据流特征构建超边,所述超边用于构建超图,所述超图用于作为网络入侵检测模型;

根据所述网络入侵检测模型检测网络入侵。

2.根据权利要求1所述的方法,其特征在于,所述方法还包括:

通过所述经过数据聚合处理的网络流量数据,得到多组数据作为训练集,预先对所述网络入侵检测模型进行训练,得到经训练后的网络入侵检测模型。

3.根据权利要求1所述的方法,其特征在于,所述基于决策树根据所述DNS解析特征以及所述数据流特征构建超边,包括:

对于提取的不同类的所述DNS解析特征以及所述数据流特征进行超边构造,分别构建决策树进行分类;

将所述决策树分类的叶子节点的所有样本划分到相同的超边之中,构建所述超边。

4.根据权利要求3所述的方法,其特征在于,所述超图用于作为网络入侵检测模型,包括:

对不同模态特征构建的超边赋予不同权值,构建超图的矩阵;

从所述超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作:

$$Y = Dv^{-\frac{1}{2}}HW\dot{D}_e^1H^TD_v^{-\frac{1}{2}}X\Theta$$

其中,X表示输入节点的特征矩阵,H为超图矩阵,D_v表示顶点度矩阵,D_e表示边的度矩阵,Θ表示学习参数;

在所述超图进行卷积运算之后,连接两个线性层作为分类器作为网络入侵检测模型。

5.根据权利要求1所述的方法,其特征在于,所述将经过数据聚合处理的网络流量数据进行特征提取,得到所述网络流量数据中的DNS解析特征以及数据流特征,包括:

将经过数据聚合处理的网络流量数据进行特征提取,得到所述DNS解析特征,所述DNS解析特征至少包括如下之一:DNS数据包中域名的语义特征、DNS数据包中的域名的注册信息、DNS数据包中的域名的注册信息、DNS数据包中的域名的解析信息、DNS通信数据包的特征。

6.根据权利要求1所述的方法,其特征在于,将经过数据聚合处理的网络流量数据进行特征提取,得到所述数据流特征,所述数据流特征至少包括如下之一:流的五元组信息、流中数据包的个数、流中数据包细节、时间特征。

7.根据权利要求1所述的方法,其特征在于,所述数据聚合处理,包括:

通过网络设备捕获原始数据包,将所述原始数据包分为DNS数据包和与通信流数据包;

根据所述原始数据包的DNS数据包中提取出请求的域名以及域名对应的IP地址;

根据所述原始数据包的通信流数据包中提取出流数据源以及目的IP地址;

如果所述通信流数据包的流数据源或目的IP地址与DNS数据包解析的IP地址相同,则将二者数据包划分到同一个分组中作为数据聚合处理的结果。

8.一种网络入侵检测装置,其特征在于,所述装置包括:

特征提取模块，用于将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征；

模型构建模块，用于基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型；

检测模块，用于根据所述网络入侵检测模型检测网络入侵。

9.一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中存储有计算机程序，其中，所述计算机程序被设置为运行时执行所述权利要求1至7任一项所述的方法。

10.一种电子装置，包括存储器和处理器，其特征在于，所述存储器中存储有计算机程序，所述处理器被设置为运行所述计算机程序以执行所述权利要求1至7任一项所述的方法。

网络入侵检测方法以及装置、存储介质、电子装置

技术领域

[0001] 本申请涉及网络安全技术接领域，具体而言，涉及一种网络入侵检测方法以及装置、存储介质、电子装置。

背景技术

[0002] 网络攻击是指攻击者通过对网络信息设备存在的漏洞和安全缺陷对系统和资源进行攻击，比如DDoS攻击、中间人劫持、网络钓鱼攻击等。

[0003] 相关技术中，网络攻击的检测方法大多对流量数据中有限的数据包进行检测，容易被攻击者绕过且由于模型学习能力有限难以检测到新类型的攻击。

[0004] 针对相关技术中网络攻击类型复杂、存在新类型攻击的问题，目前尚未提出有效的解决方案。

发明内容

[0005] 本申请的主要目的在于提供一种网络入侵检测方法以及装置、存储介质、电子装置，以解决网络攻击类型复杂、存在新类型攻击的问题。

[0006] 为了实现上述目的，根据本申请的一个方面，提供了一种网络入侵检测方法。

[0007] 根据本申请的网络入侵检测方法包括：

[0008] 将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征；

[0009] 基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型；

[0010] 根据所述网络入侵检测模型检测网络入侵。

[0011] 在一些实施例中，所述方法还包括：

[0012] 通过所述经过数据聚合处理的网络流量数据，得到多组数据作为训练集，预先对所述网络入侵检测模型进行训练，得到经训练后的网络入侵检测模型。

[0013] 在一些实施例中，所述基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，包括：

[0014] 对于提取的不同类的所述DNS解析特征以及所述数据流特征进行超边构造，分别构建决策树进行分类；

[0015] 将所述决策树分类的叶子节点的所有样本划分到相同的超边之中，构建所述超边。

[0016] 在一些实施例中，所述超图用于作为网络入侵检测模型，包括：

[0017] 对不同模态特征构建的超边赋予不同权值，构建超图的矩阵；

[0018] 从所述超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作：

[0019] $$Y = Dv^{-\frac{1}{2}} HW\dot{D}_e^1 H^T D_v^{-\frac{1}{2}} X\Theta$$

[0020] 其中，X表示输入节点的特征矩阵，H为超图矩阵，D_v表示顶点度矩阵，D_e表示边的度矩阵，Θ表示学习参数；

[0021] 在所述超图进行卷积运算之后，连接两个线性层作为分类器作为网络入侵检测模型。

[0022] 在一些实施例中，所述将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征，包括：

[0023] 将经过数据聚合处理的网络流量数据进行特征提取，得到所述DNS解析特征，所述DNS解析特征至少包括如下之一：DNS数据包中域名的语义特征、DNS数据包中的域名的注册信息、DNS数据包中的域名的注册信息、DNS数据包中的域名的解析信息、DNS通信数据包的特征。

[0024] 在一些实施例中，将经过数据聚合处理的网络流量数据进行特征提取，得到所述数据流特征，所述数据流特征至少包括如下之一：流的五元组信息、流中数据包的个数、流中数据包细节、时间特征。

[0025] 在一些实施例中，所述数据聚合处理，包括：

[0026] 通过网络设备捕获原始数据包，将所述原始数据包分为DNS数据包和与通信流数据包；

[0027] 根据所述原始数据包的DNS数据包中提取出请求的域名以及域名对应的IP地址；

[0028] 根据所述原始数据包的通信流数据包中提取出流数据源以及目的IP地址；

[0029] 如果所述通信流数据包的流数据源或目的IP地址与DNS数据包解析的IP地址相同，则将二者数据包划分到同一个分组中作为数据聚合处理的结果。

[0030] 为了实现上述目的，根据本申请的另一方面，提供了一种网络入侵检测装置。

[0031] 根据本申请的网络入侵检测装置包括：

[0032] 特征提取模块，用于将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征；

[0033] 模型构建模块，用于基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型；

[0034] 检测模块，用于根据所述网络入侵检测模型检测网络入侵。

[0035] 为了实现上述目的，根据本申请的又一方面，提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机程序，其中，所述计算机程序被设置为运行时执行所述方法。

[0036] 为了实现上述目的，根据本申请的再一方面，提供了一种电子装置，包括存储器和处理器，所述存储器中存储有计算机程序，所述处理器被设置为运行所述计算机程序以执行所述的方法。

[0037] 在本申请实施例中网络入侵检测方法以及装置、存储介质、电子装置，通过将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征，然后基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型，最后根据所述网络入侵检测模型检测网络入侵。本申请中的方法结合网络攻击中流相关的多维特征对攻击行为进行刻画，并使用了高阶关联关系的超图模型进行检测，同时使用了迁移学习思想对模型进行有

效更新。

附图说明

[0038] 构成本申请的一部分的附图用来提供对本申请的进一步理解，使得本申请的其它特征、目的和优点变得更明显。本申请的示意性实施例附图及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

[0039] 图1是根据本申请实施例的网络入侵检测方法的流程示意图；

[0040] 图2是根据本申请实施例的网络入侵检测装置的结构示意图

[0041] 图3是根据本申请实施例的网络入侵检测方法中分类器的构建原理示意图。

具体实施方式

[0042] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。

[0043] 需要说明的是，本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本申请的实施例。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0044] 在本申请中，术语“上”、“下”、“左”、“右”、“前”、“后”、“顶”、“底”、“内”、“外”、“中”、“竖直”、“水平”、“横向”、“纵向”等指示的方位或位置关系为基于附图所示的方位或位置关系。这些术语主要是为了更好地描述本申请及其实施例，并非用于限定所指示的装置、元件或组成部分必须具有特定方位，或以特定方位进行构造和操作。

[0045] 并且，上述部分术语除了可以用于表示方位或位置关系以外，还可能用于表示其他含义，例如术语“上”在某些情况下也可能用于表示某种依附关系或连接关系。对于本领域普通技术人员而言，可以根据具体情况理解这些术语在本申请中的具体含义。

[0046] 此外，术语“安装”、“设置”、“设有”、“连接”、“相连”、“套接”应做广义理解。例如，可以是固定连接，可拆卸连接，或整体式构造；可以是机械连接，或电连接；可以是直接相连，或者是通过中间媒介间接相连，又或者是两个装置、元件或组成部分之间内部的连通。对于本领域普通技术人员而言，可以根据具体情况理解上述术语在本申请中的具体含义。

[0047] 入侵检测(IDS)是网络安全的核心要素，主要目的是识别网络和计算机系统中入侵者引起的异常行为。根据如何进行入侵检测，IDS可以实现误用检测和异常检测。误用检测通过对已知的攻击行为进行分析，提取攻击特征，建立相应的攻击签名库，利用文件或者网络流量数据与攻击签名的匹配情况判断入侵行为。目前误用检测时发展最成熟，应用最广泛的技术。然而由于误用检测的基础是已知攻击签名和可获取的数据载荷包，因此误用检测通常难以应对加密攻击检测和零日攻击检测问题。

[0048] 发明人研究时发现,网络入侵检测近似的实现方案:

[0049] (1)传统的基于统计的和规则匹配的方法,通过分析统计数据来理解知识和规则。如利用协方差矩阵分析从原始数据协方差矩阵中通过矩阵分解提取规则,之后使用规则识别网络攻击。此外还有利用信息熵的方法,当攻击发生时被攻击主机IP地址出现的频率有所增加,这类方法有广义熵度量(GE),粗糙熵离群点检测等。

[0050] (2)基于机器学习的检测方法。通过人工提取数据包的特征之后使用机器学习算法进行分类,如K近邻算法(KNN),混合聚类算法,Boosting算法等。

[0051] (3)深度学习算法检测算法。这类算法对数据包头的相应字段提取特征,之后输入模型进行训练。此外有一些针对加密载荷直接输入网络进行训练,这类算法有多层感知机,卷积神经网络,循环神经网络,自动编码器等。

[0052] 不同于以上的仅仅对于流描述的方法,本申请实施例中的方法针对大规模网络环境数据进行分析,将DNS解析行为和流量通信行为进行结合,通过多种特征对攻击者行为进行描述,最后使用关联关系的超图卷积网络模型进行特征聚合和分类。

[0053] 相关技术中的数据处理方法往往关注于流量本身或者DNS解析的过程,无法体现出攻击的细节信息,而本申请实施例中的方法通过DNS解析行为、流量通信行为综合进行考虑攻击者行为并提取多维特征进行分析,可以更加详细地描述出攻击的整个过程,因此可以检测出更加隐蔽的攻击。

[0054] 此外,本申请实施例中使用的超图模型描述了不同攻击行为之间的高阶关联关系,对相关的攻击进行特征聚合可以实现高精度的检测效果。因此本申请具有检测精度高,可以检测到零日攻击的优点。

[0055] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0056] 如图1所示,该方法包括如下的步骤S101至步骤S103:

[0057] 步骤S101,将经过数据聚合处理的网络流量数据进行特征提取,得到所述网络流量数据中的DNS解析特征以及数据流特征。

[0058] 在网络环境中,提取DNS数据包中的域名和解析的IP。对解析到的IP,提取和它相关的网络流数据包,结合DNS数据包和相关的流数据包组成一个分组。使用同一个分组中不同类型的数据包构建多模态数据。

[0059] 步骤S102,基于决策树根据所述DNS解析特征以及所述数据流特征构建超边,所述超边用于构建超图,所述超图用于作为网络入侵检测模型。

[0060] 考虑DNS解析相关特征的原因在于,网络通信需要对域名进行解析,解析行为可以反应攻击特征。

[0061] 此外还需要考虑数据流相关特征,由于现有的检测方法大多对流量数据中有限的数据包进行检测,容易被攻击者绕过。

[0062] 步骤S103,根据所述网络入侵检测模型检测网络入侵。

[0063] 通过构建的超图,可用以网络入侵检测模型进行网络入侵的检测,使用超图的网络攻击检测模型。同时在超图中使用决策树的超边的构建方法,优化了超图的建立过程。

[0064] 作为本实施例中的优选,所述方法还包括:通过所述经过数据聚合处理的网络流量数据,得到多组数据作为训练集,预先对所述网络入侵检测模型进行训练,得到经训练后

的网络入侵检测模型。

[0065] 在训练时，从预处理数据中采样N个数据，构建大小为N的超图，在这N个数据中采用5折交叉验证方法，其中4份作为训练集，剩余1份作为验证集，模型训练完成后可以在线部署使用，当产生新数据时，进行实时检测。

[0066] 作为本实施例中的优选，所述基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，包括：对于提取的不同类的所述DNS解析特征以及所述数据流特征进行超边构造，分别构建决策树进行分类；将所述决策树分类的叶子节点的所有样本划分到相同的超边之中，构建所述超边。

[0067] 传统的图中一个边连接两个点。在本申请的实施例中使用了超图，一个边可以连接任意多个点。通过采用超图学习的方法，通过将多个相似实体连接在一起构建超边，超边结合在一起形成超图。之后使用超图神经网络进行卷积操作特征聚合，最后通过神经网络的分类器进行节点分类。

[0068] 对于超图分为谱分析方法，神经网络方法和一些其他方法，本申请的实施例中优选使用神经网络的方法。

[0069] 预处理环节时将经过预处理的相关的数据分组进行编码，同一组数据当作一个整体样本进行数据处理。之后特征提取环节对一组数据中提取DNS解析特征和流通信特征。接着对于提取的不同类的特征进行超边构造，分别构建决策树进行分类。这里不使用决策树分类的结果，而是把决策树分类的叶子节点的所有样本划分到相同的超边之中。在超图卷积运算之后，模型连接两个线性层作为分类器进行恶意检测。

[0070] 作为本实施例中的优选，所述超图用于作为网络入侵检测模型，包括：对不同模态特征构建的超边赋予不同权值，构建超图的矩阵；从所述超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作：

[0071] $$Y = Dv^{-\frac{1}{2}} HW\dot{D}_e^1 H^T D_v^{-\frac{1}{2}} X\Theta$$

[0072] 其中，X表示输入节点的特征矩阵，H为超图矩阵，D_v表示顶点度矩阵，D_e表示边的度矩阵，Θ表示学习参数；在所述超图进行卷积运算之后，连接两个线性层作为分类器作为网络入侵检测模型。

[0073] 由于不同类型特征对结果的影响程度是不同的，所以需要对不同模态特征构建的超边赋予不同权值。

[0074] 作为本实施例中的优选，所述将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征，包括：将经过数据聚合处理的网络流量数据进行特征提取，得到所述DNS解析特征，所述DNS解析特征至少包括如下之一：DNS数据包中域名的语义特征、DNS数据包中的域名的注册信息、DNS数据包中的域名的注册信息、DNS数据包中的域名的解析信息、DNS通信数据包的特征。

[0075] 将经过数据聚合处理的网络流量数据进行特征提取，得到所述DNS解析特征。

[0076] 所述DNS数据包中域名的语义特征，由于攻击者往往使用程序来生成大量的恶意域名，通过域名字符串可以反应此类特征，具体包括但不限于：

[0077] 域名字符串长度，表示域名字符串的个数；

[0078] 域名中数字占比，表示域名中数字个数除以域名字符串长度；

[0079] 域名的信息熵,表示用域名中每个字符出现的频率计算的熵值。

[0080] 所述DNS数据包中的域名的注册信息是指攻击者为了实现攻击往往会注册一批相似的域名来进行IP解析,具体包括但不限于:

[0081] 域名的注册人,表示注册域名的个体或者企业名称。

[0082] 域名的注册商,表示注册的域名商家。

[0083] 域名的注册时间,表示注册域名的日期。

[0084] 域名到期时间,表示域名到期的日期。

[0085] 域名的DNS服务器,表示用来解析域名的DNS服务器。

[0086] 域名的状态,表示域名解析的状态信息。

[0087] 域名的联系邮箱域名联系人的邮箱。

[0088] 联系电话,表示域名联系人的电话。其中,域名的注册时间和域名到期时间计算出年、月、日、时、分、秒六个特征,其余特征使用其相应的个数进行表示。

[0089] 所述DNS数据包中的域名的解析信息:

[0090] IP信息,表示域名解析的IP。

[0091] CNAME信息,表示域名解析的别名。

[0092] 域名的MX信息,表示域名邮件交换服务器地址。

[0093] 域名域名的默认TTL等。其中的域名域名直接使用数值,其余特征使用其相应的个数进行表示。

[0094] 所述DNS通信数据包的特征:

[0095] 时间信息,表示请求和响应包之间间隔时间。

[0096] 包头中的replycode值,表示DNS响应包中flag字段的最后4位。

[0097] 响应记录个数,表示DNS响应包中AnswerRRs字段记录个数。

[0098] 权威记录个数,表示DNS响应包中Authority RRs字段记录个数。

[0099] 附加记录个数,表示DNS响应包中Additional RRs字段记录个数。

[0100] 作为本实施例中的优选,将经过数据聚合处理的网络流量数据进行特征提取,得到所述数据流特征,所述数据流特征至少包括如下之一:流的五元组信息、流中数据包的个数、流中数据包细节、时间特征。

[0101] 将经过数据聚合处理的网络流量数据进行特征提取,得到数据流相关特征。

[0102] 所述流的五元组信息,表示即源和目的IP和端口号以及协议类型。这里对IP进行二进制编码,其余特征用数值表示。

[0103] 所述流中数据包的个数,表示相同流中在一个聚合内包含了数据包的数量。

[0104] 所述流中数据包细节,表示相同流中在一个聚合内数据包大小的最大,最小值,均值,标准差。

[0105] 所述时间特征,表示相同流中在一个聚合内数据包之间时间间隔的均值,标准差。

[0106] 作为本实施例中的优选,所述数据聚合处理,包括:通过网络设备捕获原始数据包,将所述原始数据包分为DNS数据包和与通信流数据包;根据所述原始数据包的DNS数据包中提取出请求的域名以及域名对应的IP地址;根据所述原始数据包的通信流数据包中提取出流数据源以及目的IP地址;如果所述通信流数据包的流数据源或目的IP地址与DNS数据包解析的IP地址相同,则将二者数据包划分到同一个分组中作为数据聚合处理的结果。

[0107] 从网络设备中捕获的数据往往是原始的数据包，需要对其进行相关数据包的聚合处理。访问行为往往从解析域名开始获取其IP地址，之后使用IP地址进行通信，所以把原始数据包分为DNS数据包和与通信流数据包。

[0108] 需要注意的是，这里通信流数据包泛指DNS数据包以外其他数据包。之后将相关的DNS数据包和通信流数据包聚合到相同的分组中。具体而言，

[0109] 首先，从DNS数据包中提取出请求的域名和对应的IP；

[0110] 其次，从通信流数据包中提取出源和目的IP地址；

[0111] 最后，若通信流数据包的源或目的IP地址与DNS数据包解析的IP地址相同，则把二者数据包划分到同一个分组中。聚合后的这个分组描述了整个攻击过程，作为一个统一的实体进行检测。

[0112] 由于攻击往往不会持续过长的时间，同时为了减少搜索空间的大小来减少时间复杂度。

[0113] 优选地，本申请的实施例中通过对分组中的网络流数据包数量进行限制。具体方法为在每个分组只聚合DNS解析数据包之后一段时间内的网络流数据包。

[0114] 如图2所示，本申请还提供了一种实施上述方法的装置，所述装置包括：

[0115] 特征提取模块210，用于将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征；

[0116] 模型构建模块220，用于基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型；

[0117] 检测模块230，用于根据所述网络入侵检测模型检测网络入侵。

[0118] 本申请实施例的所述特征提取模块210中在网络环境中，提取DNS数据包中的域名和解析的IP。对解析到的IP，提取和它相关的网络流数据包，结合DNS数据包和相关的流数据包组成一个分组。使用同一个分组中不同类型的数据包构建多模态数据。

[0119] 本申请实施例的所述模型构建模块220中考虑DNS解析相关特征的原因在于，网络通信需要对域名进行解析，解析行为可以反应攻击特征。

[0120] 此外还需要考虑数据流相关特征，由于现有的检测方法大多对流量数据中有限的数据包进行检测，容易被攻击者绕过。

[0121] 本申请实施例的所述模型构建模块230中通过构建的超图，可用以网络入侵检测模型进行网络入侵的检测，使用超图的网络攻击检测模型。同时在超图中使用决策树的超边的构建方法，优化了超图的建立过程。

[0122] 显然，本领域的技术人员应该明白，上述的本申请的各模块或各步骤可以用通用的计算装置来实现，它们可以集中在单个的计算装置上，或者分布在多个计算装置所组成的网络上，可选地，它们可以用计算装置可执行的程序代码来实现，从而，可以将它们存储在存储装置中由计算装置来执行，或者将它们分别制作成各个集成电路模块，或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样，本申请不限制于任何特定的硬件和软件结合。

[0123] 为了更好的理解上述网络入侵检测方法中分类器的构建的流程，以下结合优选实施例对上述技术方案进行解释说明，但不用于限定本发明实施例的技术方案。

[0124] 从网络设备中捕获的数据往往是原始的数据包，这里需要对其进行相关数据包的

聚合处理。访问行为往往从解析域名开始获取其IP地址，之后使用IP地址进行通信，所以把原始数据包分为DNS数据包和与通信流数据包。这里通信流数据包泛指DNS数据包以外其他数据包。之后把相关的DNS数据包和通信流数据包聚合到相同的分组中得到聚合数据。

[0125] 预处理环节时，把相关的数据分组进行编码，同一组数据当作一个整体样本进行数据处理。之后特征提取环节对一组数据中提取DNS解析特征和流通信特征。接着对于提取的不同类的特征进行超边构造，分别构建决策树进行分类。这里不使用决策树分类的结果，而是把决策树分类的叶子节点的所有样本划分到相同的超边之中。

[0126] 由于不同类型特征对结果的影响程度是不同的，这里对不同模态特征构建的超边赋予不同权值。之后从超图的矩阵中提取拉普拉斯矩阵并定义预设卷积操作，在超图卷积运算之后，模型连接两个线性层作为分类器进行恶意检测。

[0127] 以上所述仅为本申请的优选实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

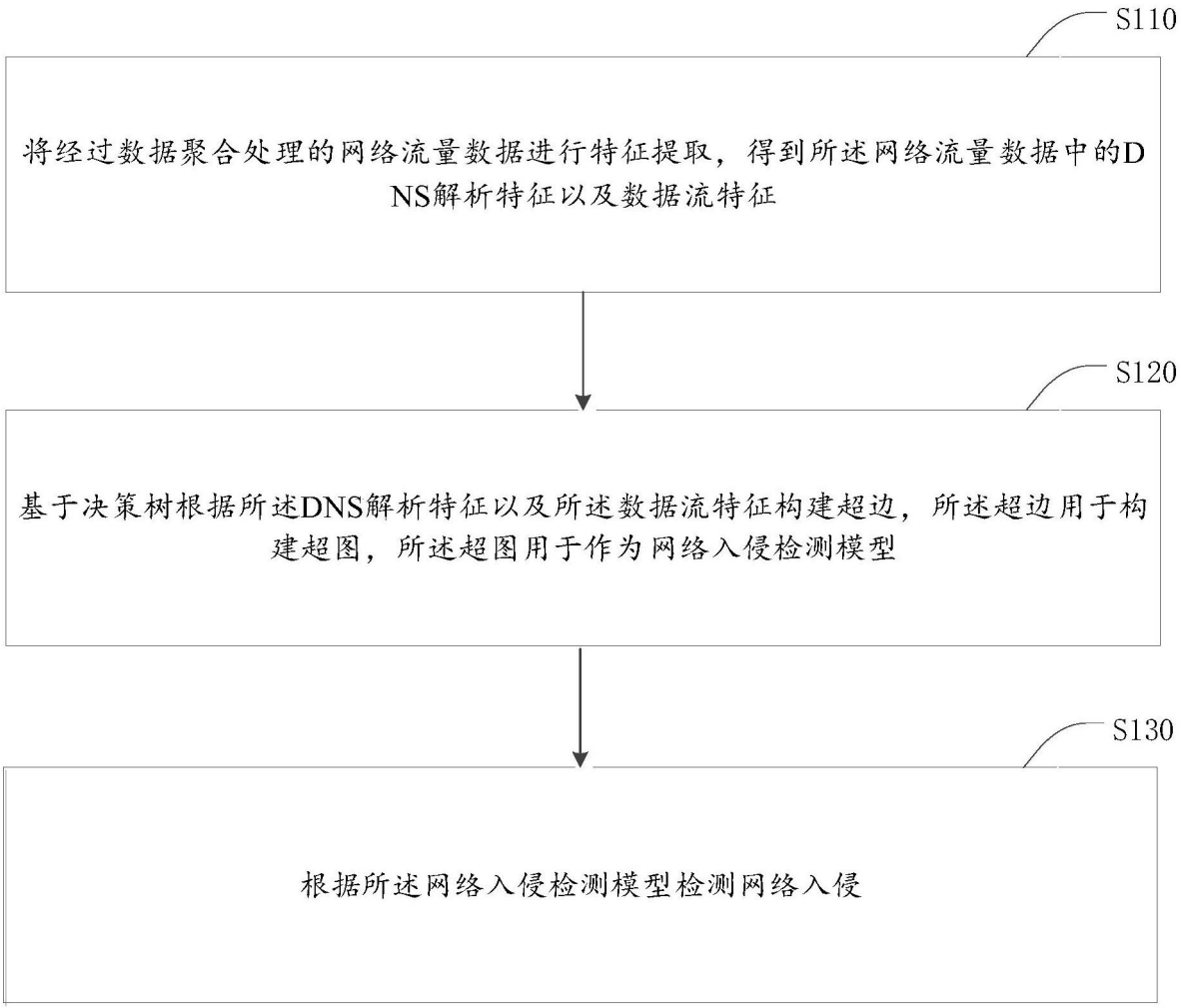

S110
将经过数据聚合处理的网络流量数据进行特征提取，得到所述网络流量数据中的DNS解析特征以及数据流特征
S120
基于决策树根据所述DNS解析特征以及所述数据流特征构建超边，所述超边用于构建超图，所述超图用于作为网络入侵检测模型
S130
根据所述网络入侵检测模型检测网络入侵

图1


200
特征提取模块210
模型构建模块220
检测模块230
网络入侵检测装置

图2

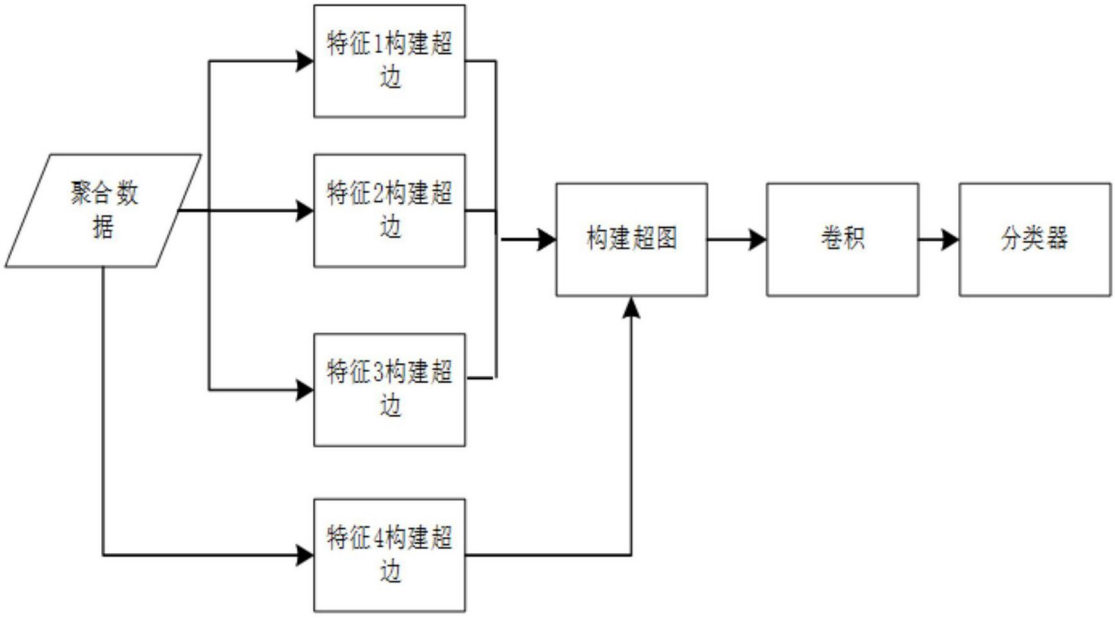

聚合数据
特征1构建超边
特征2构建超边
特征3构建超边
特征4构建超边
构建超图
卷积
分类器

图3