(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113079167 A
(43)申请公布日 2021.07.06

(21)申请号 202110389696.1

(22)申请日 2021.04.12

(71)申请人 西北工业大学
地址 710072 陕西省西安市友谊西路127号

(72)发明人 孙文韬 吴诒轩 聂来森 宁兆龙

(74)专利代理机构 北京高沃律师事务所 11569
代理人 王爱涛

(51)Int.Cl.
H04L 29/06(2006.01)
G06N 20/00(2019.01)

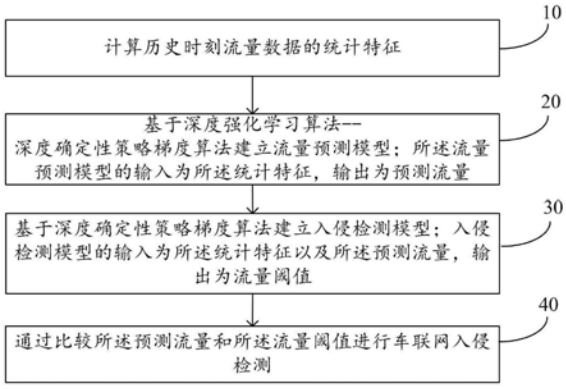
权利要求书2页 说明书7页 附图3页

(54)发明名称

一种基于深度强化学习的车联网入侵检测方法及系统

(57)摘要

本发明公开了一种基于深度强化学习的车联网入侵检测方法及系统，该方法包括：计算历史时刻流量数据的统计特征；基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型；所述流量预测模型的输入为所述统计特征，输出为预测流量；基于深度确定性策略梯度算法建立入侵检测模型；所述入侵检测模型的输入为所述统计特征以及所述预测流量，输出为流量阈值；通过比较所述预测流量和所述流量阈值进行车联网入侵检测。本发明能够兼顾车联网的复杂多变性、基础设施服务器端计算资源有限性和网络入侵检测的准确性，更符合实际。对于车联网这种复杂的系统，本发明提出的基于深度强化学习的入侵检测方法相比于目前多用的其他方法有更好的性能。



CN 113079167 A

1.一种基于深度强化学习的车联网入侵检测方法,其特征在于,包括:

步骤10:计算历史时刻流量数据的统计特征;

步骤20:基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;所述流量预测模型的输入为所述统计特征,输出为预测流量;

步骤30:基于深度确定性策略梯度算法建立入侵检测模型;所述入侵检测模型的输入为所述统计特征以及所述预测流量,输出为流量阈值;

步骤40:通过比较所述预测流量和所述流量阈值进行车联网入侵检测。

2.根据权利要求1所述的基于深度强化学习的车联网入侵检测方法,其特征在于,所述计算历史时刻流量数据的统计特征,具体包括:

步骤101:将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量;

步骤102:分别计算历史时刻基于不同协议流量的统计特征;所述统计特征包括均值、方差和稀疏性。

3.根据权利要求2所述的基于深度强化学习的车联网入侵检测方法,其特征在于,基于TCP协议的流量服从高斯分布,基于UDP协议的流量服从泊松分布。

4.根据权利要求1所述的基于深度强化学习的车联网入侵检测方法,其特征在于,所述基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;具体包括:

步骤201:建立动作-在线网络μ_1、动作-目标网络μ_1'、评价-在线网络Q_1和评价-目标网络Q_1',建立记忆库R_1;

步骤202:获取当前状态$s_{t1}=\{D(t)\}$,D(t)表示统计特征的集合;检验当前状态是否到达目标状态,到达目标状态则结束,否则由μ_1选择一个动作$a_{t1}=\{X'(t)\}$,X'(t)为流量预测值;根据选择动作到达下一状态;到达新的状态入侵检测模型会获得即时奖励值r_{t1},将转移元组$(s_{t1},a_{t1},r_{t1},s_{t1+1})$存储到$R_1$;

步骤203:从R_1中随机采样n个转移元组,建立损失函数更新评价在线网络Q_1和更新动作在线网络μ_1,然后软更新目标网络μ_1'和Q_1';

步骤204:将下一状态设置为当前状态,重复步骤步骤202。

步骤205:利用当前时刻流量数据,完成流量预测得到流量预测值X'(t)。

5.根据权利要求4所述的基于深度强化学习的车联网入侵检测方法,其特征在于,所述基于深度确定性策略梯度算法建立入侵检测模型,具体包括:

步骤301:建立动作-在线网络μ_2、动作-目标网络μ_2'、评价-在线网络Q_2和评价-目标网络Q_2',建立记忆库R_2;

步骤302:获取当前状态$s_{t2}=\{D(t),X'(t)\}$,检验当前状态是否到达目标状态,到达目标状态则结束;否则由μ_2选择一个动作$a_{t2}=\{x(t)\}$,x(t)为动态阈值。根据该动作到达下一状态;到达新的状态入侵检测模型会获得即时奖励值r_{t2},将转移元组$(s_{t2},a_{t2},r_{t2},s_{t2+1})$存储到$R_2$;

步骤303:从R_2中随机采样n个转移元组,建立损失函数更新评价在线网络Q_2和更新动作在线网络μ_2,然后软更新目标网络μ_2'和Q_2';

步骤304:将下一状态设置为当前状态,重复步骤步骤302。

6.根据权利要求1所述的基于深度强化学习的车联网入侵检测方法,其特征在于,所述通过比较所述预测流量和所述流量阈值进行车联网入侵检测,具体包括:

当预测流量大于流量阈值时,判定服务器受到DDoS攻击;

当预测流量小于流量阈值时,判定服务器处于正常状态。

7.一种基于深度强化学习的车联网入侵检测系统,其特征在于,包括:

统计特征计算模块,用于计算历史时刻流量数据的统计特征;

流量预测模型建立模块,用于基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;所述流量预测模型的输入为所述统计特征,输出为预测流量;

入侵检测模型建立模块,用于基于深度确定性策略梯度算法建立入侵检测模型;所述入侵检测模型的输入为所述统计特征以及所述预测流量,输出为流量阈值;

入侵检测模块,用于通过比较所述预测流量和所述流量阈值进行车联网入侵检测。

8.根据权利要求7所述的基于深度强化学习的车联网入侵检测系统,其特征在于,所述统计特征计算模块具体包括:

区分单元,用于将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量;基于TCP协议的流量服从高斯分布,基于UDP协议的流量服从泊松分布;

计算单元,用于分别计算历史时刻基于不同协议流量的统计特征;所述统计特征包括均值、方差和稀疏性。

一种基于深度强化学习的车联网入侵检测方法及系统

技术领域

[0001] 本发明涉及网络安全技术领域,特别是涉及一种基于深度强化学习的车联网入侵检测方法及系统。

背景技术

[0002] 随着通信网络技术的不断发展和成功应用,人们对于车联网服务提出了更高的要求,这直接导致车联网结构越来越复杂。而随着现代车辆的复杂性、连通性的不断提高,车联网的网络安全风险也变得越来越突出。为了保证网络的安全和正常运行,实时可靠的安全性增强方法是必不可少的。入侵检测系统作为一种轻量级的安全性增强方法,能够很好地侦测网络内部的和外部的威胁,并且拥有较好的成本效益与高度兼容性,目前被认为是增强车联网安全性的不错选择。

[0003] 近年来,车联网的安全性获得了广泛的关注,为了实现高效的入侵检测,提出了很多方法。基于异常的入侵检测是目前较为常用的入侵检测方法,它根据统计行为建模,对正常行为进行分析,将与正常行为的一定偏差标记为异常。随着人工智能技术的兴起,利用机器学习方法实现对正常行为的建模,对正常行为进行分析,进而标记异常行为的方法,大大提高了入侵检测的准确性与可靠性。Miao等人针对异常检测问题制定了分布式在线支持向量机,并获得了分散成本函数。Garg等人提出了一种基于受限玻尔兹曼机和支持向量机的集成方法。有学者提出利用深度神经网络可以对输入数据进行逐层学习,从而可以获得高层次的特征数据并进一步用于分类等任务。Garg等人提出了一种利用灰狼优化和卷积神经网络进行网络异常检测的混合数据处理模型。

[0004] 虽然提出了很多方法提高入侵检测的准确性,一个显著的问题在于难以选择有效的特征对正常行为建模。因此,本发明提出了一种基于深度强化学习与流量预测的入侵检测算法,可以准确可靠的实现入侵检测。

发明内容

[0005] 针对现有技术存在的问题,本发明提供了一种基于深度强化学习的车联网入侵检测方法及系统,准确可靠的实现入侵检测。

[0006] 为实现上述目的,本发明提供了如下方案:

[0007] 一种基于深度强化学习的车联网入侵检测方法,包括:

[0008] 步骤10:计算历史时刻流量数据的统计特征;

[0009] 步骤20:基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;所述流量预测模型的输入为所述统计特征,输出为预测流量;

[0010] 步骤30:基于深度确定性策略梯度算法建立入侵检测模型;所述入侵检测模型的输入为所述统计特征以及所述预测流量,输出为流量阈值;

[0011] 步骤40:通过比较所述预测流量和所述流量阈值进行车联网入侵检测。

[0012] 可选地,所述计算历史时刻流量数据的统计特征,具体包括:

[0013] 步骤101:将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量;

[0014] 步骤102:分别计算历史时刻基于不同协议流量的统计特征;所述统计特征包括均值、方差和稀疏性。

[0015] 可选地,基于TCP协议的流量服从高斯分布,基于UDP协议的流量服从泊松分布。

[0016] 可选地,所述基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;具体包括:

[0017] 步骤201:建立动作-在线网络μ_1、动作-目标网络μ_1'、评价-在线网络Q_1和评价-目标网络Q_1',建立记忆库R_1;

[0018] 步骤202:获取当前状态$s_{t1}=\{D(t)\}$,D(t)表示统计特征的集合;检验当前状态是否到达目标状态,到达目标状态则结束,否则由μ_1选择一个动作$a_{t1}=\{X'(t)\}$,X'(t)为流量预测值;根据选择动作到达下一状态;到达新的状态入侵检测模型会获得即时奖励值r_{t1},将转移元组$(s_{t1},a_{t1},r_{t1},s_{t1+1})$存储到$R_1$;

[0019] 步骤203:从R_1中随机采样n个转移元组,建立损失函数更新评价在线网络Q_1和更新动作在线网络μ_1,然后软更新目标网络μ_1'和Q_1';

[0020] 步骤204:将下一状态设置为当前状态,重复步骤步骤202。

[0021] 步骤205:利用当前时刻流量数据,完成流量预测得到流量预测值X'(t)。

[0022] 可选地,所述基于深度确定性策略梯度算法建立入侵检测模型;所述入侵检测模型的输入为所述统计特征以及所述预测流量,输出为流量阈值,具体包括:

[0023] 步骤301:建立动作-在线网络μ_2、动作-目标网络μ_2'、评价-在线网络Q_2和评价-目标网络Q_2',建立记忆库R_2;

[0024] 步骤302:获取当前状态$s_{t2}=\{D(t),X'(t)\}$,检验当前状态是否到达目标状态,到达目标状态则结束;否则由μ_2选择一个动作$a_{t2}=\{x(t)\}$,x(t)为动态阈值。根据该动作到达下一状态;到达新的状态入侵检测模型会获得即时奖励值r_{t2},将转移元组$(s_{t2},a_{t2},r_{t2},s_{t2+1})$存储到$R_2$;

[0025] 步骤303:从R_2中随机采样n个转移元组,建立损失函数更新评价在线网络Q_2和更新动作在线网络μ_2,然后软更新目标网络μ_2'和Q_2';

[0026] 步骤304:将下一状态设置为当前状态,重复步骤步骤302。

[0027] 可选地,所述通过比较所述预测流量和所述流量阈值进行车联网入侵检测,具体包括:

[0028] 当预测流量大于流量阈值时,判定服务器受到DDoS攻击;

[0029] 当预测流量小于流量阈值时,判定服务器处于正常状态。

[0030] 本发明还提供了一种基于深度强化学习的车联网入侵检测系统,包括:

[0031] 统计特征计算模块,用于计算历史时刻流量数据的统计特征;

[0032] 流量预测模型建立模块,用于基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型;所述流量预测模型的输入为所述统计特征,输出为预测流量;

[0033] 入侵检测模型建立模块,用于基于深度确定性策略梯度算法建立入侵检测模型;所述入侵检测模型的输入为所述统计特征以及所述预测流量,输出为流量阈值;

[0034] 入侵检测模块,用于通过比较所述预测流量和所述流量阈值进行车联网入侵检测。

[0035] 可选地，所述统计特征计算模块具体包括：

[0036] 区分单元，用于将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量；基于TCP协议的流量服从高斯分布，基于UDP协议的流量服从泊松分布；

[0037] 计算单元，用于分别计算历史时刻基于不同协议流量的统计特征；所述统计特征包括均值、方差和稀疏性。

[0038] 根据本发明提供的具体实施例，本发明公开了以下技术效果：

[0039] 本发明提拱了一种基于深度强化学习的车联网入侵检测方法及系统，该方法包括：计算历史时刻流量数据的统计特征；基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型；所述流量预测模型的输入为所述统计特征，输出为预测流量；基于深度确定性策略梯度算法建立入侵检测模型；所述入侵检测模型的输入为所述统计特征以及所述预测流量，输出为流量阈值；通过比较所述预测流量和所述流量阈值进行车联网入侵检测。本发明能够兼顾车联网的复杂多变性、基础设施服务器端计算资源有限性和网络入侵检测的准确性，更符合实际。对于车联网这种复杂的系统，本发明提出的基于深度强化学习的入侵检测方法相比于目前多用的其他方法有更好的性能。

附图说明

[0040] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0041] 图1为本发明实施例基于深度强化学习的车联网入侵检测方法的流程图；

[0042] 图2为本发明的应用场景；

[0043] 图3为本发明采用的深度强化学习原理图；

[0044] 图4为本发明应用于CICDDoS2019数据集时在时间相对误差上与已有办法比较；

[0045] 图5为本发明应用于CICDDoS2019数据集时在时间相对误差的累积分布函数上与已有办法比较；

[0046] 图6为本发明应用于CICDDoS2019数据集时在偏差上与已有办法比较。

具体实施方式

[0047] 下面将结合本发明实施例中的附图，对本发明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发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，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，都属于本发明保护的范围。

[0048] 为使本发明的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0049] 如图1所示，一种基于深度强化学习的车联网入侵检测方法包括以下步骤：

[0050] 步骤10：计算历史时刻流量数据的统计特征。具体包括：

[0051] 步骤101：将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量；

[0052] 步骤102：分别计算历史时刻基于不同协议流量的统计特征；所述统计特征包括均

值、方差和稀疏性。基于TCP协议的流量服从高斯分布，基于UDP协议的流量服从泊松分布。

[0053] 步骤20：基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型；所述流量预测模型的输入为所述统计特征，输出为预测流量。具体包括：

[0054] 步骤201：建立动作-在线网络μ_1、动作-目标网络μ_1'、评价-在线网络Q_1和评价-目标网络Q_1'，建立记忆库R_1；

[0055] 步骤202：获取当前状态$s_{t1}=\{D(t)\}$，D(t)表示统计特征的集合；检验当前状态是否到达目标状态，到达目标状态则结束，否则由μ_1选择一个动作$a_{t1}=\{X'(t)\}$，X'(t)为流量预测值；根据选择动作到达下一状态；到达新的状态入侵检测模型会获得即时奖励值r_{t1}，将转移元组$(s_{t1},a_{t1},r_{t1},s_{t1+1})$存储到$R_1$；

[0056] 步骤203：从R_1中随机采样n个转移元组，建立损失函数更新评价在线网络Q_1和更新动作在线网络μ_1，然后软更新目标网络μ_1'和Q_1'；

[0057] 步骤204：将下一状态设置为当前状态，重复步骤步骤202。

[0058] 步骤205：利用当前时刻流量数据，完成流量预测得到流量预测值X'(t)。

[0059] 步骤30：基于深度确定性策略梯度算法建立入侵检测模型；所述入侵检测模型的输入为所述统计特征以及所述预测流量，输出为流量阈值。具体包括：

[0060] 步骤301：建立动作-在线网络μ_2、动作-目标网络μ_2'、评价-在线网络Q_2和评价-目标网络Q_2'，建立记忆库R_2；

[0061] 步骤302：获取当前状态$s_{t2}=\{D(t),X'(t)\}$，检验当前状态是否到达目标状态，到达目标状态则结束；否则由μ_2选择一个动作$a_{t2}=\{x(t)\}$，x(t)为动态阈值。根据该动作到达下一状态；到达新的状态入侵检测模型会获得即时奖励值r_{t2}，将转移元组$(s_{t2},a_{t2},r_{t2},s_{t2+1})$存储到$R_2$；

[0062] 步骤303：从R_2中随机采样n个转移元组，建立损失函数更新评价在线网络Q_2和更新动作在线网络μ_2，然后软更新目标网络μ_2'和Q_2'；

[0063] 步骤304：将下一状态设置为当前状态，重复步骤步骤302。

[0064] 步骤40：通过比较所述预测流量和所述流量阈值进行车联网入侵检测。当预测流量大于流量阈值时，判定服务器受到DDoS攻击；当预测流量小于流量阈值时，判定服务器处于正常状态。

[0065] 具体实施例：

[0066] 如图2所示，攻击者在基础设施服务器端施加DDoS攻击，我们在基础设施服务器端获取流量数据并计算统计特征，然后反馈给服务器管理者，服务器管理者通过深度确定性策略梯度方法(如图3所示)计算当前时刻流量的预测值并进而完成对自身的入侵检测，实现提高自身安全性的目的。

[0067] 步骤一：在基础设施服务器端，收集流量数据。将过去100个时刻的流量数据(X(t-100),...,X(t-1))，区分为基于TCP协议的流量数据(T(t-100),...,T(t-1))与基于UDP协议的流量数据(U(t-100),...,U(t-1))。分别计算过往100个时刻的基于不同协议流量的均值、方差和稀疏性等统计特征。其中稀疏性为过往100个时刻内非零流量的个数，统计特征的集合记作D(t)。

[0068] 步骤二：在基础设施服务器端，利用流量数据实现流量预测。利用过往时刻流量的统计特征，基于深度强化学习算法--深度确定性策略梯度算法对流量变化行为进行建模。

在模型收敛之后，对于给定的状态s_t^P模型会生成一个全局最优解。

[0069] 具体包括如下步骤：

[0070] 步骤A：建立动作-在线网络μ_1、动作-目标网络μ_1'、评价-在线网络Q_1和评价-目标网络Q_1'，建立记忆库R_1。

[0071] 步骤B：获取当前状态s_{t1}＝{D(t)}，D(t)表示统计特征的集合；检验当前状态是否到达目标状态，到达目标状态则结束，否则由μ_1选择一个动作a_{t1}＝{X'(t)}，X'(t)为流量预测值；根据选择动作到达下一状态；到达新的状态入侵检测模型会获得即时奖励值r_{t1}，将转移元组$(s_{t1},a_{t1},r_{t1},s_{t1+1})$存储到$R_1$。奖励值计算公式如下：$r_{t1}=\frac{\lambda^P}{D_M(X(t),X'(t))}$，其中$\lambda^P=\begin{cases}1, & X'(t)>0\\ 0.5, & X'(t)<0\end{cases}$。

[0072] 步骤C：从R_1中随机采样64个转移元组，取学习率γ为0.9，建立损失函数更新评价在线网络Q_1和更新动作在线网络μ_1，然后软更新目标网络μ_1'和Q_1'。

[0073] 损失函数如下：$L(\theta^{Q_1})=\frac{1}{64}\sum_t^{64}(y_{t1}-Q_1(s_{t_1},a_{t1},\theta^{Q_1}))^2$，其中$y_{t1}=r_{t1}+0.9Q_1'(s_{t1+1},\mu_1'(s_{t1+1},\theta^{\mu_1'}),\theta^{Q_1'})$，θ为网络参数，动作在线网络μ的策略梯度如下：$\nabla_{\theta^{\mu_1}}J(\mu_1)=\frac{1}{64}\sum_t^{64}(\nabla_a Q(s,a,\theta^{Q_1})|_{s=s_{t1},a=\mu_1(s_{t1})}\times\nabla_{\theta^{\mu}}\mu_1(s,\theta^{\mu})|_{s=s_{t1}})$，软更新方法如下：取软更新系数τ为0.01，$\theta^{Q_1'}\leftarrow 0.01*\theta^{Q_1}+0.99*\theta^{Q_1'},\theta^{\mu'}\leftarrow 0.01*\theta^{\mu}+0.99*\theta^{\mu'}$。

[0074] 步骤D：将下一状态设置为当前状态，重复步骤B。

[0075] 步骤E：利用步骤一收集到当前时刻所需数据，完成流量预测得到流量预测值X'(t)。

[0076] 步骤三：在基础设施服务器端，利用流量数据与流量预测结果建立入侵检测模型。

[0077] 首先建立入侵检测模型：$P=\begin{cases}1, & X'(t)>x\\ 0, & X'(t)<x\end{cases}$，当预测流量值大于阈值x时，判定服务器受到DDoS攻击，小于阈值x时，则认为服务器处于正常状态。然后利用步骤二中过往时刻的流量特征D(t)以及流量预测的结果X'(t)，基于深度确定性策略梯度算法生成阈值x。

[0078] 步骤A：建立动作-在线网络μ_2、动作-目标网络μ_2'、评价-在线网络Q_2和评价-目标网络Q_2'，建立记忆库R_2。

[0079] 步骤B获取当前状态s_{t2}＝{D(t)，X'(t)}，检验当前状态是否到达目标状态，到达目标状态则结束；否则由μ_2选择一个动作a_{t2}＝{x(t)}，x(t)为动态阈值。根据该动作到达下一状态；到达新的状态入侵检测模型会获得即时奖励值r_{t2}，将转移元组$(s_{t2},a_{t2},r_{t2},s_{t2+1})$存储到$R_2$。奖励值计算公式如下：$r_{t2}$＝i·f(abs(X(t)-x))。其中，i与当前阈值是否判定攻击成功有关，成功为1，失败为-1，$f(x)=\begin{cases}l, & x>l\\ x, & x\leqslant l\end{cases}$，奖励值范围l取10，abs(·)表示绝对值。

[0080] 步骤C：从R_2中随机采样64个转移元组，取学习率γ为0.85，据此建立损失函数更

新评价在线网络Q_2和更新动作在线网络μ_2，然后软更新目标网络μ_2'和Q_2'。损失函数如下：$L(\theta^{Q_2})=\frac{1}{64}\sum_{t}^{64}(y_{t2}-Q_2(s_{t2},a_{t2},\theta^{Q_2}))^2$，其中$y_{t2}=r_{t2}+0.85*Q_2'(s_{t2+1},\mu_2'(s_{t2+1},\theta^{\mu_2'}),\theta^{Q_2'})$，θ为网络参数，动作在线网络μ的策略梯度如下：$\nabla_{\theta^{\mu 2}}J(\mu_2)=\frac{1}{64}\sum_{t}^{64}(\nabla_a Q(s,a,\theta^Q)|_{s=s_{t2},a=\mu_2(s_{t2})}\times\nabla_{\theta^{\mu 2}}\mu_2(s,\theta^{\mu_2})|_{s=s_{t2}})$，软更新方法如下：取软更新系数τ为0.01，$\theta^{Q_2'}\leftarrow 0.01*\theta^{Q_2}+0.99*\theta^{Q_2'},\theta^{\mu_2'}\leftarrow 0.01*\theta^{\mu_2}+0.99*\theta^{\mu_2'}$。

[0081] 步骤D：将下一状态设置为当前状态，重复步骤B。

[0082] 步骤四：服务器根据入侵检测模型检测自身是否受到入侵，如收到入侵则启动防护措施。

[0083] 步骤A：根据步骤一收集数据，根据步骤二和步骤三计算流量预测结果与入侵检测模型。

[0084] 步骤B：根据入侵检测模型生成当前时刻的阈值x(t)，根据阈值x(t)与当前时刻流量值X(t)判断自身是否受到入侵。如受到入侵，报警并启动防护措施，否则重复步骤A。

[0085] 尽管现在已有多种方法对车联网网络安全进行了研究，但是本发明提出的方法兼顾车联网的复杂多变性、基础设施服务器端计算资源有限性和网络入侵检测的准确性，所提方法更符合实际。车联网本身存在时变性和不确定性导致网络是一个复杂的大系统，数学模型的复杂性和精确性往往难以满足网络的实时需求。基础设施端的服务器的计算资源十分有限，而大多准确性高的方法对计算资源有着较高的要求。深度强化学习方法不依赖于数学模型和先验知识，通过不断的试错和与环境的不断交互获得知识，具有自学习的能力。而结合深度学习的特点使其不必保存大量的状态空间，节省了服务器的大量存储空间，并且对于计算资源的要求与运行速率均优于寻常的深度学习方法。因此，对于车联网这种复杂的系统，本发明提出的基于深度强化学习的入侵检测方法相比于目前多用的其他方法有更好的性能。

[0086] 本发明将研究内容应用于CICDDoS2019数据集，并将得到的结果与已有方法比较，研究本发明提出的基于深度强化学习的车联网入侵检测方法是否具有优越性以及普适性。

[0087] 本发明首先研究了实验步骤中流量预测方法的准确性，并分别与已有的方法进行了对比实验与分析。多分形小波模型通过利用流量的自相似性预测未来流量的趋势，稀疏正则矩阵分解利用流量的时空特征，即相邻元素的值更为接近这一特征来实现流量预测。本发明将所提方法得出的结果与其比较，以检验本方法的优越性。图4与图5展示了三种方法在时间相对误差上的表现，图6展示了三种方法在偏差上的表现。如图4、图5所示，本发明所提出方法在时间相对误差上均低于其他两种方法，这说明本发明所提算法拥有更强的对于流量变化趋势的预测能力。但是，如图6所示，本发明所提方法在偏差方面大于其他两种方法，这是因为模型对于预测非零真实流量值时会获得更高的奖励，这会使模型对于非零值的预测更加准确，但也同时导致了对全局流量偏差的增大。考虑模型的最终目的是把握流量的变化趋势，从而实现入侵检测，这样的结果是相对可以接受的。

[0088] 然后本发明研究了实验步骤中入侵检测方法的准确性，并分别与已有的方法进行了对比实验与分析。主成分分析法与稀疏正则矩阵分解法均是通过算法特性实现流量重构或预测。本发明将所提方法得出的结果与其比较，结果如表1所示，本发明所提方法在准确率、召回率和F1值方面均高于其他两种方法。这表明本发明所提方法不仅能够快速准确的

完成入侵检测，在流量预测方面也有较好的表现，十分具有实际应用意义。

[0089] 表1

[0090]

指标	DDPG	SRMF	PCA
TPR	100%	92.48%	3.47%
FPR	1.14%	1.12%	1.26%
Precision	99.69%	99.68%	93.96%
F1	0.9984	0.9594	0.0669

[0091] 此外，本发明还提供了一种基于深度强化学习的车联网入侵检测系统，包括：

[0092] 统计特征计算模块，用于计算历史时刻流量数据的统计特征。

[0093] 流量预测模型建立模块，用于基于深度强化学习算法--深度确定性策略梯度算法建立流量预测模型；所述流量预测模型的输入为所述统计特征，输出为预测流量。

[0094] 入侵检测模型建立模块，用于基于深度确定性策略梯度算法建立入侵检测模型；所述入侵检测模型的输入为所述统计特征以及所述预测流量，输出为流量阈值。

[0095] 入侵检测模块，用于通过比较所述预测流量和所述流量阈值进行车联网入侵检测。

[0096] 其中，所述统计特征计算模块具体包括：

[0097] 区分单元，用于将服务器接收的流量区分为基于TCP协议与基于UDP协议的流量；基于TCP协议的流量服从高斯分布，基于UDP协议的流量服从泊松分布。

[0098] 计算单元，用于分别计算历史时刻基于不同协议流量的统计特征；所述统计特征包括均值、方差和稀疏性。

[0099] 本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言，由于其与实施例公开的方法相对应，所以描述的比较简单，相关之处参见方法部分说明即可。

[0100] 本文中应用了具体个例对本发明的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本发明的方法及其核心思想；同时，对于本领域的一般技术人员，依据本发明的思想，在具体实施方式及应用范围上均会有改变之处。综上所述，本说明书内容不应理解为对本发明的限制。

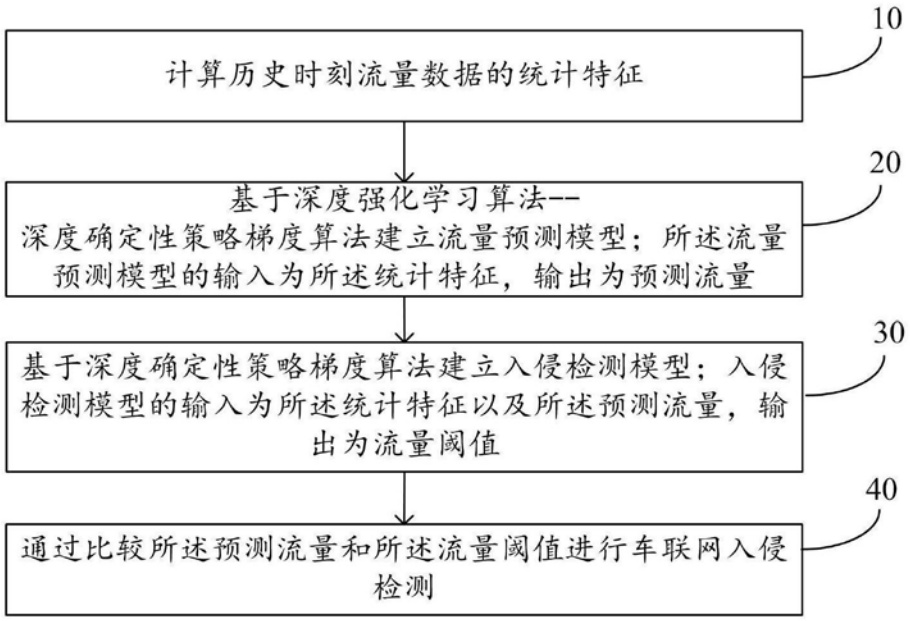

10
计算历史时刻流量数据的统计特征
20
基于深度强化学习算法--
深度确定性策略梯度算法建立流量预测模型；所述流量预测模型的输入为所述统计特征，输出为预测流量
30
基于深度确定性策略梯度算法建立入侵检测模型；入侵检测模型的输入为所述统计特征以及所述预测流量，输出为流量阈值
40
通过比较所述预测流量和所述流量阈值进行车联网入侵检测

图1

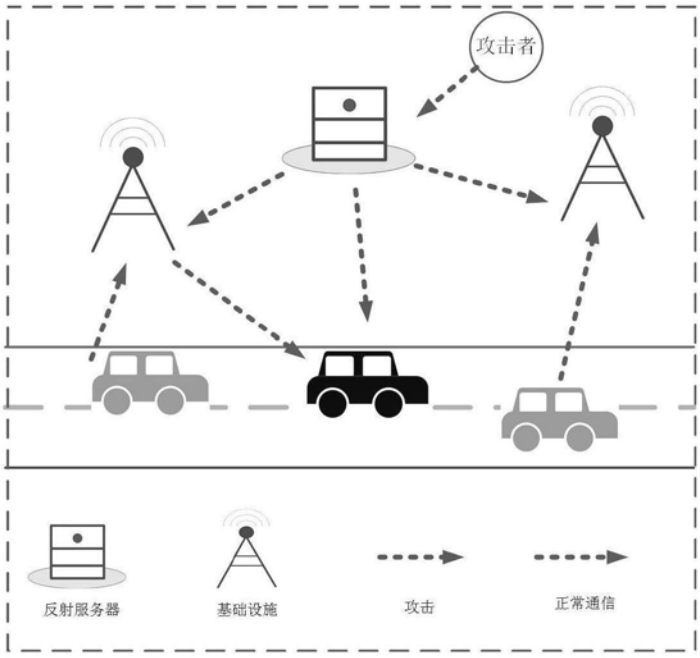

攻击者
反射服务器
基础设施
攻击
正常通信

图2

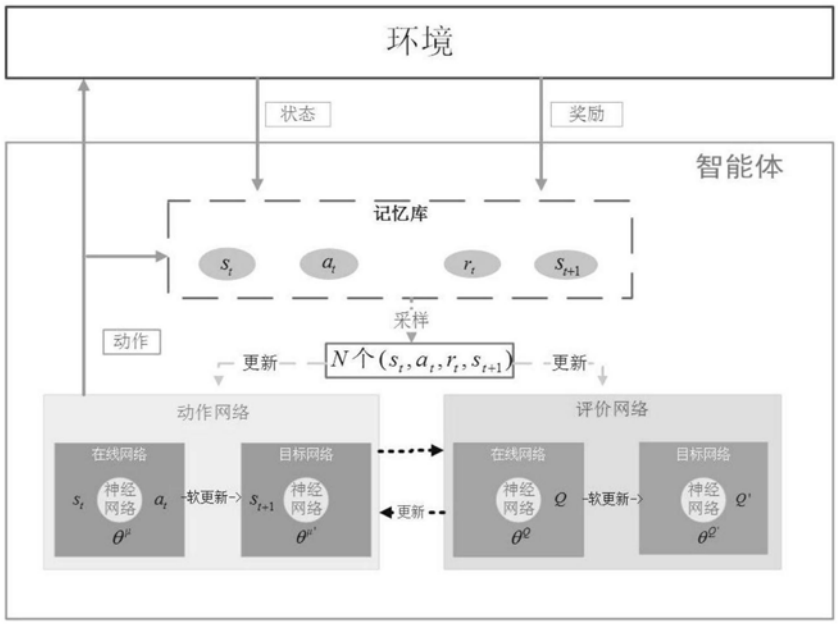

环境
状态
奖励
智能体
记忆库
s_t
a_t
r_t
s_{t+1}
采样
动作
更新
N个(s_t, a_t, r_t, s_{t+1})
更新
动作网络
评价网络
在线网络
目标网络
在线网络
目标网络
神经网络
s_t
a_t
软更新
s_{t+1}
神经网络
更新
神经网络
Q
软更新
神经网络
Q'
θ^{μ}
$\theta^{\mu'}$
θ^{Q}
$\theta^{Q'}$

图3

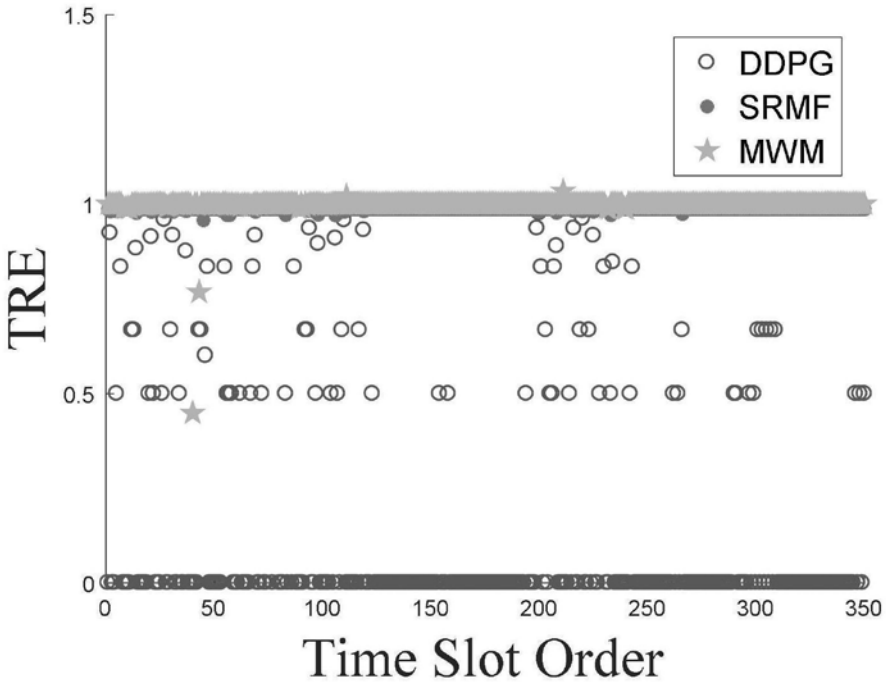

DDPG
SRMF
MWM
TRE
1.5
1
0.5
0
0
50
100
150
200
250
300
350
Time Slot Order

图4

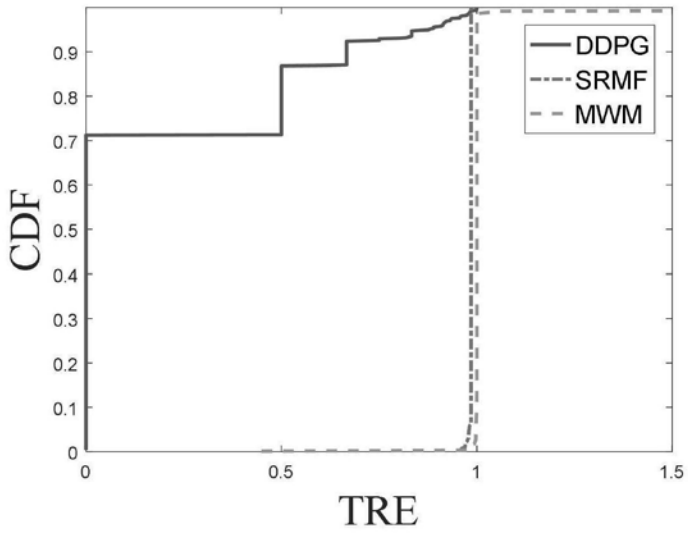

DDPG
SRMF
MWM
CDF
TRE
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
0.5
1
1.5

图5

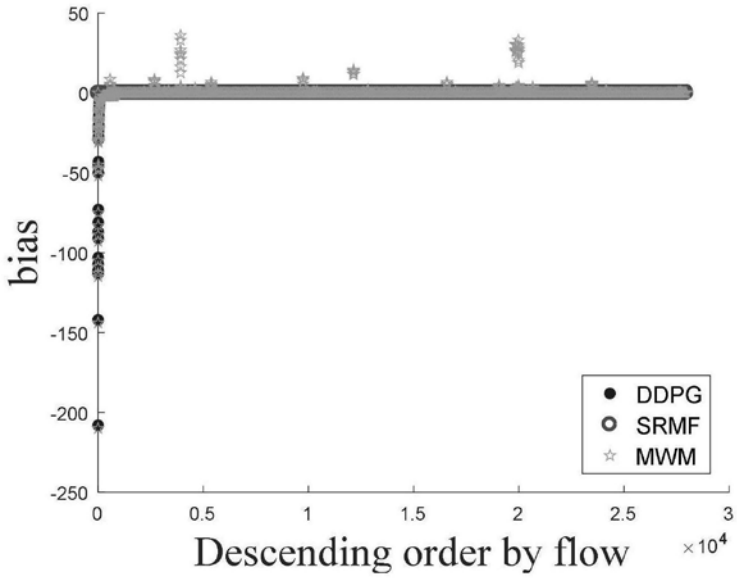

DDPG
SRMF
MWM
bias
Descending order by flow
×10^4
50
0
-50
-100
-150
-200
-250
0
0.5
1
1.5
2
2.5
3

图6