Computer Science and Application 计算机科学与应用, 2020, 10(12), 2177-2188
Published Online December 2020 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2020.1012229

Hans汉斯

基于流量统计的车联网Sybil攻击行为检测

吕佳健[1,2]，马　进[1,2*]，陈秀真[1,2]

[1]上海交通大学网络空间安全学院，上海
[2]上海市信息安全综合管理技术研究重点实验室，上海
Email: lvjiajian@sjtu.edu.cn, *majin@sjtu.edu.cn

收稿日期：2020年11月24日；录用日期：2020年12月9日；发布日期：2020年12月16日

摘　要

在传统的VANETS (Vehicle Ad-hoc Networks)向IoV (Internet of Vehicles)转变的过程中，车联网范围不断扩大以及车联网内流量不断增加，使得传统入侵检测系统难以识别攻击信息或者无法及时给出反馈。Sybil攻击作为车联网环境中的一种重要攻击类型，通过模拟正常车辆的运行特性造成道路信息混乱、阻塞，为车辆的安全驾驶带来巨大挑战。本文面向车联网环境中的sybil攻击行为，使用基于流量统计的入侵检测分析方法来对sybil攻击行为进行鉴别。在数据收集方面，利用仿真工具Veins模拟实际车联网环境中的信息传输过程；在分类器选用方面，利用机器学习中的MLP (Multilayer Perceptron)神经网络，训练能够识别sybil攻击的入侵检测模型；在可视化方面，设计具有便捷交互性的用户界面来进行展示和预警。检测实验表明，本文所设计的入侵检测系统在区分正常流量和sybil攻击行为上平均达到了85%以上的准确率，同时能够在可视化界面上及时地反馈预测结果。

关键词

车联网，入侵检测，流量统计，sybil攻击，机器学习

Sybil Attack Behavior Detection for V2x Based on Traffic Statistics

Jiajian Lv[1,2], Jin Ma[1,2*], Xiuzhen Chen[1,2]

[1]SJTU School of Cyber Science and Engineering, Shanghai
[2]Shanghai Key Laboratory of Integrated Administration Technologies for Information Security, Shanghai
Email: lvjiajian@sjtu.edu.cn, *majin@sjtu.edu.cn

Received: Nov. 24th, 2020; accepted: Dec. 9th, 2020; published: Dec. 16th, 2020

*通讯作者。

文章引用: 吕佳健, 马进, 陈秀真. 基于流量统计的车联网 Sybil 攻击行为检测[J]. 计算机科学与应用, 2020, 10(12): 2177-2188. DOI: 10.12677/csa.2020.1012229

Abstract

In the process of transforming traditional VANETS (Vehicle Ad-hoc Networks) to IoV (Internet of Vehicles), the scope of the Internet of Vehicles continues to expand and the traffic within the Internet of Vehicles continues to increase, making it difficult for traditional intrusion detection systems to identify attack information or provide timely information Feedback. As an important type of attack in the Internet of Vehicles environment, Sybil attacks cause confusion and obstruction of road information by simulating the operating characteristics of normal vehicles, which brings huge challenges to the safe driving of vehicles. This article is oriented to Sybil attacks in the Internet of Vehicles environment, and uses traffic statistics-based intrusion detection and analysis methods to identify Sybil attacks. In terms of data collection, the simulation tool Veins is used to simulate the information transmission process in the actual car networking environment; in terms of classifier selection, the use of MLP (Multilayer Perceptron) neural network in machine learning is used to train an intrusion detection model that can identify Sybil attacks; in terms of visualization, we design a convenient and interactive user interface for display and warning. Detection experiments show that the intrusion detection system designed in this paper achieves an average accuracy of over 85% in distinguishing between normal traffic and sybil attacks, and can provide timely feedback on the prediction results on the visual interface.

Keywords

Internet of Vehicles, Intrusion Detection, Traffic Statistics, Sybil Attack, Machine Learning

1. 引言

随着互联网汽车产业的快速发展以及人工智能、无人驾驶技术的日趋成熟，新生的物联网领域正在驱动传统的 VANETS 向 IoV 转变。在转变的过程中车联网的安全问题也日益突出。为保障车联网安全，国内外的专家学者提出了多种方案，主要包括：1) 设计一款安全路由协议，2) 为各个攻击建立威胁模型，3) 设计实现入侵检测系统。本文针对车联网领域的入侵检测系统做进一步的研究。车联网相对于传统互联网来说，对计算能力、存储能力、安全性等要求更高，而传统的车联网中的车载入侵检测模型存在诸多问题，包括：

a) 车载设备性能差距较大。入侵检测系统需要保证各个车辆的兼容性，这对当前复杂繁多的车载设备来说，实现的难度较大。

b) 车载设备资源受限。车载终端的计算与存储能力都较弱，使得入侵检测模型所能利用的资源受限，难以加快检测效率。

c) 车联网环境复杂。入侵检测模型在预测时需要保证低时延，复杂的车联网环境中需要鉴别更多的流量信息，加重车载入侵检测模型的负担。

d) 车载设备本身的安全威胁。入侵检测模型本身所运行的环境受到很大的挑战，如 CAN 总线等均可能存在安全漏洞，有被攻击的风险。

同时，车联网中存在很多的攻击技术，如 sybil 攻击以其模拟正常车辆运动特性的方式伪造数据包，造成道路信息混乱阻塞，从而导致车流拥堵和交通事故等问题，为车辆的安全驾驶和复杂的车联网环境带来了更多的安全挑战。

本文充分考虑车联网相较于传统互联网的特殊性，采用基于流量统计的方式，利用当前在诸多领域取得突破的机器学习技术来实现能够识别 sybil 攻击的入侵检测系统：

a) 针对车载端性能差异大的特点，本文使用了一个全局的入侵检测系统，使得车载设备不再作为检测入侵行为的载体，而作为数据的收集终端。

b) 从鉴别 sybil 攻击的方式来看，流量统计的方法恰好契合恶意节点发送攻击信息的方式，从全局的角度收集车联网内部的流量信息，再通过利用 MLP 建立的入侵检测模型判断是否发生 sybil 攻击。从实验结果看来，这一方法在区分正常流量和 sybil 攻击行为上平均达到了 85%以上的准确率。

2. Sybil 攻击行为及其检测技术

Sybil 攻击可以归类为 V2X (Vehicle to Everything)中最危险的和最难检测的攻击行为之一[1]。在 sybil 攻击情形中，模拟车辆可以假装具有多个身份。换句话说，网络中的其他节点无法区分信息是来自一辆车辆还是来自多于一辆车辆。攻击者的主要目的是根据其目标来塑造网络。例如，攻击者可以操纵其他车辆的行为，使其他车辆偏离预定路线。除了是最危险的攻击形式之一之外，sybil 攻击也是最难发现的攻击之一。在使用地理路由的网络上，sybil 攻击更具风险，因为攻击者通过发送有关其位置的错误信息来声称车辆处于多个位置。此外，它还可以模拟发生在实际位置以外位置的事件。本小节将介绍 sybil 攻击的主要分类以及常用的 sybil 检测技术。

2.1. Sybil 攻击分类

车辆使用 IEEE 802.11p 网络(也称为 ITS-G5 网络)进行通信。所有车辆均使用 5.9 GHZ 频段定期广播 V2X 消息。每条消息都包含车辆的假名(临时身份)和一些运动学信息(位置、速度、前进方向等)。ITS (Intelligent Traffic System)中 PKI (Public Key Infrastructure)向车辆提供一个长期证书和几个短期证书，称为假名证书。这些证书用于签署 V2X 消息。车辆经常更改其假名以避免跟踪并保护其隐私。每辆车在特定时间段内使用单个假名证书进行签名。

为了确保车辆连续发送 V2X 消息的能力，必须同时提供几个有效的假名给每个车辆。欧盟委员会建议最多使用 100 个有效假名证书。当 ITS 的可用假名不足时，它将请求发送到 PKI，用新证书重新填充证书池。按照规定，车辆在一定时间内不能使用多个假名证书来签署其消息。但是行为异常的车辆可能有意同时使用多个有效的假名证书，以此模拟车联网中的幽灵节点(不存在但能被车联网中其他车辆感知到的模拟节点)，这会导致 sybil 攻击。根据攻击者的目标，sybil 攻击可能采用不同的形式。参考 Joseph 等人的文献[2]将其分为 4 类：

1) 随机数据填充拒绝服务 sybil 攻击(Dos Random Sybil Attack)

如图 1 所示，攻击者使用随机数据创建消息，消息包含车辆的假名和运动学信息(位置、速度、前进方向等)。攻击者对每个发送的消息使用不同的假名，并将此种消息发送到车联网环境中，使得正常车辆接收到此类信息时，无法及时鉴定消息真伪，容易引发事故。

2) 数据回送 sybil 攻击(Data Replay Sybil Attack)

如图 2 所示，此攻击误导入侵检测系统将合法车辆报告为恶意车辆。攻击者选择受害车辆，在接收到受害者车辆的消息后，创建同样包含受害车辆运动学的消息，发送到车联网环境中，使得在其他正常节点看来，同一时空位置具有两辆不同的车辆，且无法判断哪个为正常车辆，造成混乱。

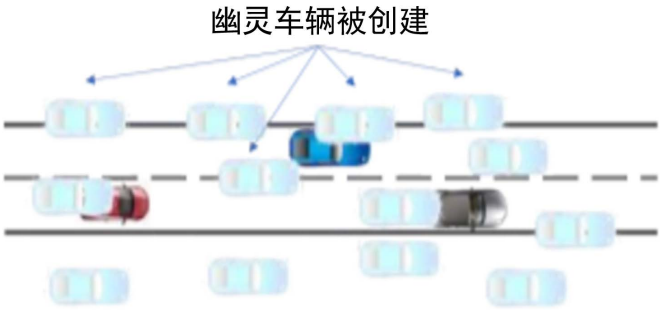

Figure 1. Dos random sybil attack
图 1. 随机数据填充拒绝服务 sybil 攻击

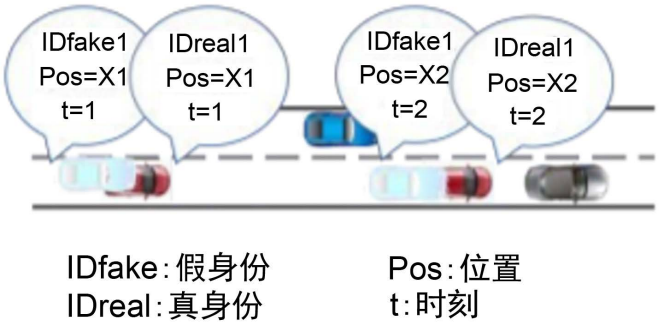

Figure 2. Data replay sybil attack
图 2. 数据回送 sybil 攻击

3) 破坏性拒绝服务 sybil 攻击(Dos Disruptive Sybil Attack)

此攻击是 Dos Random Sybil Attack 和 Data Replay Sybil Attack 的组合。如图 3 所示，攻击者在每个消息中使用不同的假名，但不会用随机数据填充它们。相反，传输的消息是基于从相邻的正常节点接收到的数据。此种攻击和 Data Replay Sybil Attack 之间的区别在于它没有跟随一名受害节点，攻击者收集数据信息，在不定的时间内发送一连串数据包来模拟一个正常的车辆行驶流程。

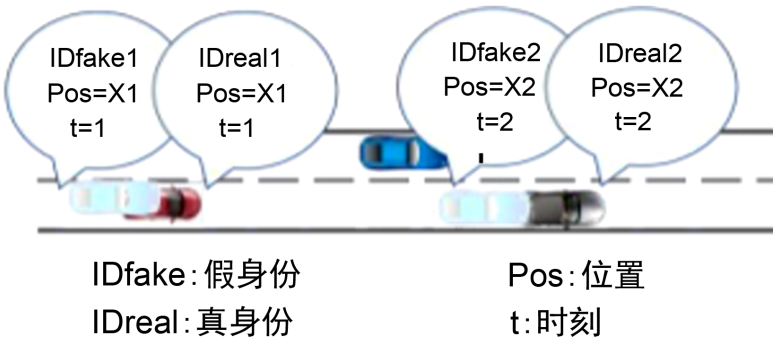

Figure 3. Dos disruptive sybil attack
图 3. 破坏性拒绝服务 sybil 攻击

4) 交通拥堵 sybil 攻击(Grid Sybil Attack)

如图 4 所示，攻击者使用有效的假名来模拟多辆幽灵车辆。使得恶意车辆通信范围内的车辆接收到虚假消息，并得出道路上发生拥堵的结论。攻击者可以智能地计算出幽灵车的运动数据，从而使假消息具有合理性和连贯性。

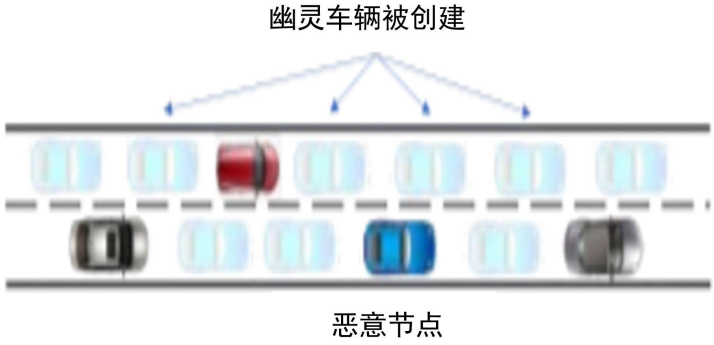

Figure 4. Grid sybil attack
图 4. 交通拥堵 sybil 攻击

2.2. 常用的 sybil 攻击检测技术

Pouyan 等提出了三种检测本地 Sybil 攻击的方法[3]：1) 资源测试方法，假定无线网络实体不能同时在同一信道上发送和接收。这种检测方法在车载网络中无效，因为攻击者可能具有多个发送和接收消息的渠道；2) 位置验证方法，假定车辆只能同时定位在一个位置，但在车联网中正常节点广播的数据包可以进行一定程度的伪造，因此位置验证方法存在一定的缺陷，即无法判断幽灵节点是否为正常节点；3) 基于加密和身份验证的方法,假定使用 PKI 足以检测不具备密钥的消息，但是具有有效密钥材料的合法实体也可以执行 sybil 攻击。

Ghaleb 等提出了一种基于人工神经网络的局部不良行为检测模型[4]，使用一些特征来确定车辆是否具有异常行为。但是仅利用从单一车辆捕获的信息进行本地检测是不够的，例如 Grid Sybil Attack 模拟的幽灵节点均符合正常车辆的运行特征。因此需要一个可以访问更多不良行为报告的全局系统。

本文提出一个全局的检测系统来收集车联网内一段时间的流量信息，即通过流量统计的方式，判断局域网内发生的 sybil 攻击行为，并识别出恶意车辆，对恶意车辆进行监控和预警。

2.3. 基于流量统计的 sybil 入侵检测

由于车联网的车辆间信息交互的特点，车联网中每一个拥有身份标识的车辆均相当于一个无线网点，其发送的消息或者其本身均能够被车联网中的其他车辆所感知。所以当一个车联网内的恶意节点在传输攻击信息时，无法指定这个攻击信息是否直接发送给某个特定的车辆，也就是说，攻击行为是不具备定向性的。在恶意节点周围的正常节点均能够接收到此攻击消息，因此，sybil 攻击通常会使得整个车联网陷入混乱。而流量统计的方式从多个车辆角度来识别恶意节点，恰好契合攻击行为的不定向性，从全局的角度收集车联网内部的流量信息，通过入侵检测模型判断是否发生 sybil 攻击，再对识别为 sybil 攻击类型的信息进行统计，以此识别出恶意车辆。

3. 基于流量统计的入侵检测系统设计

3.1. 入侵检测系统架构和流程

基于流量统计的车联网入侵检测系统整体架构由以下三个模块组成：数据收集模块、分类器模块和可视化模块，如图 5 所示。

1) 数据收集模块包含路边单元 RSU 和车联网中的正常节点。该模块收集车联网环境中车辆之间传递的消息信息，形成报文，并以报文的形式发送给分类器模块；

2) 分类器模块接收到数据后对数据进行预处理，并使用已经训练好的分类器进行分类判断，并将分

类结果进行保存，同时保存发送此数据包的车辆身份标识；在此过程中收集到的车辆标识信息和发送的数据包信息均保存在全局入侵检测系统的数据模块中；

3) 可视化模块通过对数据模块内容进行提取和总结，显示进行恶意攻击的车辆信息和攻击类型，并且对识别的准确率加以统计，对各个攻击的预警级别也进行判断，统计的结果显示在可视化界面中。

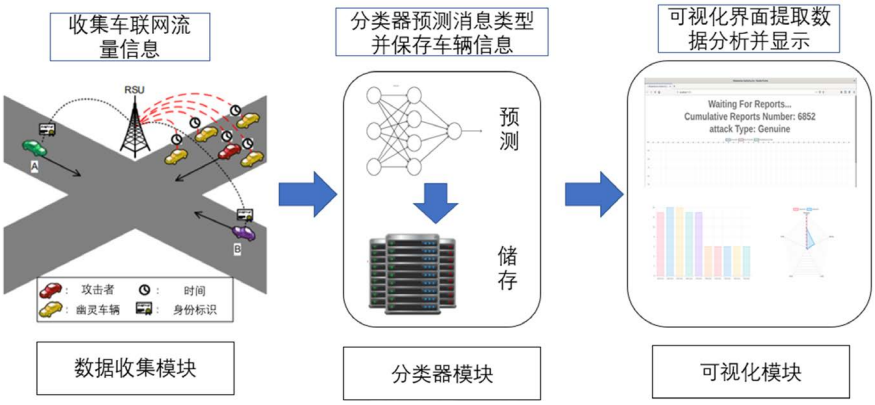

Figure 5. The overall architecture of the car network intrusion detection system
图 5. 车联网入侵检测系统整体架构

基于流量统计的车联网入侵检测流程如图 6 所示，主要包括数据获取、数据处理、机器学习入侵检测和上报四个过程：首先收集车联网流量，进行流量特征置信度计算；再将收集到的流量信息内容进行归一化、去重、下采样等数据预处理操作，得到较好的训练数据集和测试数据集，并以此来训练一个基于 MLP 的入侵检测模型，用于识别攻击类型；最后将检测结果保存处理，用于可视化模块的显示功能。

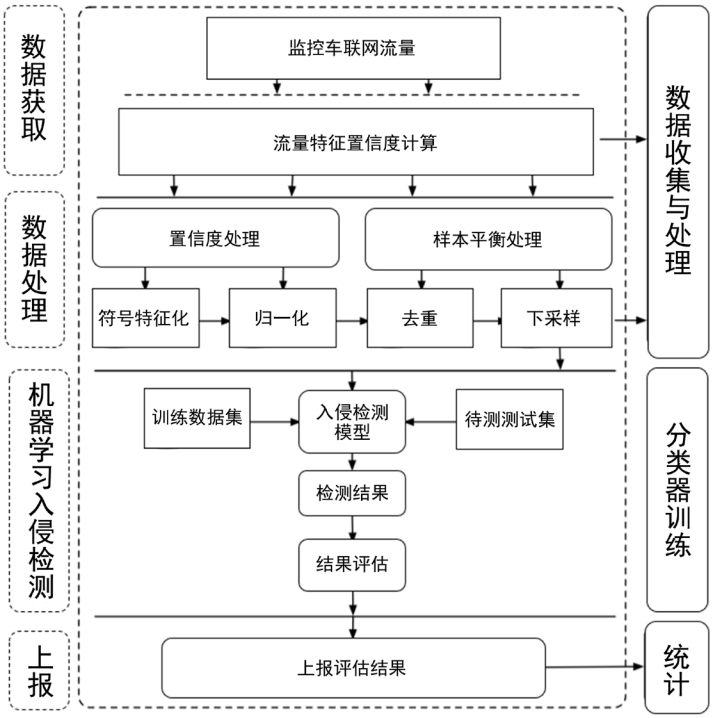

Figure 6. Flow chart of intrusion detection
图 6. 入侵检测流程图

3.2. 代码设计

入侵检测系统主要函数说明如表 1 所示。

Table 1. Class and main function description
表 1. 类与主要函数说明

类名称	函数名称	函数解释
MAMain	maMain(reportJsonString)	处理从 veins 中接收到的消息，提取特征并经过入侵检测模型检测，得到结果传递给数据库模块保存，是数据处理到转入储存的中心逻辑部分。
	readBSM(reportJsonStirng)	根据流量提取范围特征和合理性特征，传递给训练好的入侵检测模型，得到预测结果并返回给 maMain()。
MAStorge	pingma()	可视化模块提取数据的接口，将数据传递给可视化模块用于实现可视化。
	addreport(msgContent, result)	保存每一个车联网消息内容和预测结果，以及发送此消息的车辆的身份标识信息。
	get_Msg()	提取数据储存模块内的数据，反馈给服务器。
Masever	do_get()	接收数据收集模块收集的车联网流量，并将收集到的数据作为 MaMain()流程的输入。
	do_post()	通过请求的方式提取入侵检测模型的判断结果、车辆标识等信息，用于 web 可视化界面。

4. 特征概述和分类器训练

4.1. Sybil 攻击特征提取

采用 Veins 仿真框架模拟车联网中车辆间的信息传递，收集流量信息，将此流量信息作为整个实验过程中的训练数据和测试数据。Veins 中的道路交通仿真由 SUMO 执行，而网络仿真由 OMNeT++与物理层建模工具包 MiXiM 一起执行，可以为无线电干扰以及静态和移动障碍物的遮挡提供精确的模型，可以模拟现实环境中信号传递因物理阻碍和干扰带来的影响。Veins 还提供了数据传递 API，使得入侵检测系统可以通过端口的方式实时接收车联网中的模拟数据。Veins 规定：模拟车辆中的正常节点在模拟环境中运行的各个运动学信息必须遵从表 2 中的特征范围。

Table 2. Feature range
表 2. 特征范围

模拟特征	描述
Position Range (位置范围)	车辆所处的位置需要合理，Veins 规定车辆位置必须限定在模拟道路上。
Speed Range (速度范围)	车辆的速度要在合理的范围内，Veins 规定模拟环境中车辆的最大行驶速度不超过 40 km/h。
Heading Range (方向范围)	车辆行驶的方向在一定时间内需要保持一致。
Acceleration Range (加速度范围)	车辆的加速度要在合理的范围内，Veins 规定模拟环境中车辆的最大加速度不超过 3 m/s^2。
Maximum Separation Distance (最大接收距离)	Veins 规定车辆接收消息的最大距离为 420 m。
Beacon Frequency (发送消息频率)	Veins 规定车辆发送消息的频率不高于 0.99 个/秒。

模拟车联网环境中的车辆各个运动特征均应该保持在上述范围之内。2.1 小节所提到的各种 sybil 攻

击方法中，恶意节点发送攻击类型的信息，其消息内容或是具有一定的随机性，或是和车联网中其他模拟车辆的消息内容重叠。因此，利用车联网环境的运动学范围特征可以对消息内容的置信度进行评估，并以此作为判断 sybil 攻击的依据。

4.1.1. Sybil 攻击特征概述

从 2.1 节中 sybil 攻击的攻击方式来看，攻击消息主要和 Position、Speed、Heading 和 Acceleration 四个特征有很大的关联。然而，某一时刻的车辆运动学信息在可信范围之内并不代表下一时刻该车辆的运动学信息也在可信范围之内，比如在 Grid Sybil Attack 中，恶意节点可以通过细致的计算来模拟正常车辆的运动，使得在一段时间内幽灵车辆的运动信息符合正常车辆运动信息，因此，仅通过上述四个特征难以判断。同时，车辆运动学信息的各个方面在可信范围之内并不代表整个消息是可信的，比如，在接收到信息的两个时刻，同一车辆的位置和速度都可信，但车辆以可信范围内的速度运行，经过两个时刻的时间间隔，并不足以从前一个时刻位置达到后一时刻位置。因此，我们还需要考虑特征之间的相互关联性，以此来增强车辆消息可信度的判断依据。

4.1.2. Sybil 攻击特征分析

考虑车辆运动学特征的可信范围和运动学特征之间的联系，本次实验利用了 Joseph 等人提出的置信度计算方法 CaTch [5]，该解决方案在执行基本的不良行为检查时会考虑到移动性信息的不确定性，再进一步展开运动学基本特征，并根据各个基本特征之间的联系计算置信度。总结如下：

1) Range plausibility (rP) ：接收范围合理性，代表接收消息车辆和发送消息车辆距离小于模拟特征规定的最大距离 420m。

2) Position plausibility (pP)：位置合理性，车辆的位置限定在道路上，而不能出现在障碍物中。

3) Speed plausibility (sP)：速度合理性，模拟车辆的速度不能超过规定的最大速度。

4) Position consistency (pC)：位置一致性，同一模拟车辆的两个时刻的位置距离小于一定的阈值。

5) Speed consistency (sC)：速度一致性，两个时刻内同一车辆的速度变化需要小于加速度的最大阈值。

6) Position speed consistency (psC)：位置速度一致性，两个不同时刻同一车辆的位置和速度之间存在关联，位置的变化必须保持速度在阈值之内。

7) Positon speed memory consistency(psmC)：多时刻位置速度一致性，记录了多个时刻同一车辆的位置和速度之间的关联，位置变化与速度变化保持关联。

8) Position heading consistency (phC)：位置方向一致性，两个时刻的同一车辆的位置和车辆行进的方向之间存在关联。车辆在一段时间内的运动和其方向保持一致，和位置变化矢量之间存在关联。

9) Beacon frequency (bF)：信标频率，模拟车辆发送消息具有一定的时间间隔，发送消息频率不超过 0.99 个/秒。

10) Intersection check (inT)：交叉检测，同一时刻同一位置不能存在多个车辆。

11) Sudden appearance (sA)：新出现车辆的检测，接收到的新出现的车辆位置必须在合理的范围内。

12) Kalman Filter Tracking：使用卡尔曼过滤方式[6]，通过计算得到几种检测特征：

-卡尔曼位置速度一致性(kPSCP，kPSCS，kPSCSP，kPSCSS)，

-卡尔曼位置一致性(kPC)，

-卡尔曼位置加速度一致性(kPAC)，

-卡尔曼速度一致性(kSC)。

根据表 2 总结的本次实验的 6 个特征范围，可以计算得到范围特征、合理性特征总计 18 个特征的置信度作为本次实验的检测特征，用于判断该消息是否为攻击类型的消息。

4.2. 基于 MLP 的分类器训练

4.2.1. 数据集收集

本文使用 F2MD 框架来评估基于流量统计的解决方案[7]。Paris Scalay 是 Veins 内置的由 SUMO 生成的模拟场景，我们将 Paris Scalay 交通方案用于车辆跟踪。模拟场景的拓扑结构由市区组成，住宅区道路和主干道与高速公路相连。在模拟环境中可以投放发送恶意攻击(sybil 攻击)的节点，并通过监听端口的方式收集车联网模拟环境内部的流量信息。在每一种场景中，恶意节点的比例为 5%。我们分以下两种情况收集流量数据。

1) 600 s 时长的、包含单一 sybil 攻击的流量：对于 2.1 小节中提到的各种 sybil 攻击类型，我们分别收集了 600 s 的流量统计结果，通过数据收集模块的整理和计算，得到了如表 3 的结果。

Table 3. Sample information in 600 s traffic
表 3. 600 s 流量中的样本信息

流量包含的攻击类型	样本总数/条	正负样本比值
DosRandomSybilAttack	50,770	6.97:1
DataReplaySybilAttack	43,279	7.02:1
DosDisruptiveSybilAttack	53,238	6.04:1
Grid Sybil Attack	54,663	6.42:1

2) 1200 s 时长的、包含各种 sybil 攻击的流量：经过数据收集模块的整理和计算，并去除重复数据后，得到样本总数为 88,645。各类型样本比例为 5.03:1.09:1:1.34:1.22。

收集 600 s 时长的、包含单一 sybil 攻击的流量数据，其意义在于判断各种 sybil 攻击类型消息和正常流量消息间的区别，同时判断各个特征对于识别攻击的贡献程度。而收集 1200 s 时长的、包含各种 sybil 攻击的流量数据，意义在于训练一个多分类器来区分 sybil 攻击的类别。

4.2.2. 数据预处理

通过分析收集到的流量数据，可以总结出以下几个特点，并根据特点进行对应的数据预处理操作。

1) 数据归一化：归一化的目的就是使得预处理的数据被限定在一定的范围内，从而消除奇异样本数据导致的不良影响；同时，如若不进行归一化处理，在进行梯度下降的时候，梯度的方向就会偏离最小值的方向，走很多弯路，导致训练时间过长。

在 4.1.2 小节中，根据置信度的计算，各个特征的取值范围存在一定的差异，为了减轻这种情况带来的影响，方便后续的数据处理过程，同时加快网络学习速度，可以对各个特征进行归一化，使得所有样本的输入信号均方差相比很小。

2) 平衡正负样本：根据 4.2.1 小节统计得到的样本比例，可以观察到正负样本不够平衡，而样本不平衡会使得我们的分类模型存在严重偏向，即将攻击类型消息判断为正常消息。从测试结果上来看，会表现为有太多的 False Positive。因此，本文采用 Random Under Sample 的欠采样方式来平衡正负样本。

4.2.3. 分类器训练和结果分析

我们分别对预处理后的 600 s 样本数据进行了训练，并利用 F1 得分(F1 Score，F1 Score 是准确率和召回率整合在一起的判定标准，是统计学中用来衡量二分类模型精确度的一种指标)来衡量入侵检测模型的表现，综合考虑入侵检测模型的检测准确率和误报率，F_1 得分的计算公式如(1)所示。

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{1}$$

经过多次训练得到 F_1 得分最高的 MLP 的各项重要参数和结果平均值如表 4 所示，F_1 得分的计算公式如公式。

Table 4. Classifier parameters and results in single sybil attack detection
表 4. 单一 sybil 攻击检测分类器的参数和结果

输入信号个数	隐藏层节点	激活函数	学习率	迭代次数
18	18	Relu	0.001	200
Sybil 攻击类型	准确率	召回率	F1 得分	支持数据数量
DosRandom	0.995	0.982	0.991	4226
Data Replay	0.951	0.920	0.936	2119
Dos Disruptive	0.981	0.986	0.984	4239
Grid	0.843	0.842	0.843	2130

这一步主要是比较了各种类别的 sybil 攻击和正常消息，判断各类 sybil 攻击消息和正常消息是否具有比较大的区别。从训练结果可以看出，对于各种类型的 sybil 攻击方式，使用 MLP 进行训练和判断均达到比较好的预测效果。原因在于 Dos Random Sybil Attack 由于内容的随意填充，其大部分特征(尤其是 pP，sP 特征)的置信度很低，容易区分；而 Data Replay Sybil Attack 通过跟踪车辆的方式发送消息，其 inT 置信度数值比较低，也能被区分；Dos Disruptive Sybil Attack 因为同时具有 Dos Random 和 Data Replay 类型的特征，使得 MLP 能很好地做出检测。

Grid Sybil Attack 通过严密的计算来模拟幽灵车辆，因而相比于其他类型而言识别率较低。通过图 7 各个特征的得分情况可以看到，MLP 在判断 Grid 类型时主要考虑了卡尔曼滤波器求得的特征和 psC、psmC 等一致性特征，也验证了一致性特征在检验 sybil 攻击方面的有效性。

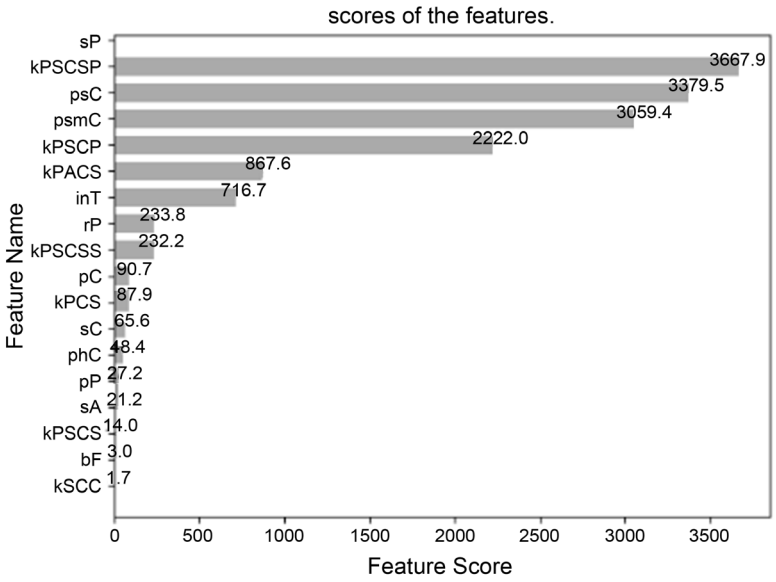

Figure 7. Two classifier feature score
图 7. 二分类器特征得分

对预处理后的 1200 s 样本数据进行二分类训练和多分类训练，经过多次测试得到 F1 分数最高的 MLP 的各项重要参数如表 5 所示。

Table 5. Two classifier parameters and results
表 5. 二分类器参数和结果

输入信号个数	隐藏层节点	激活函数	学习率	迭代次数
18	18	Relu	0.001	200
类型	准确率	召回率	F1 得分	支持数据数量
Genuine	0.861	0.954	0.903	5851
Misbehavior	0.942	0.855	0.897	5839

结合 600 s 训练集得到的结果看来，正常消息和 sybil 攻击消息的区别较为明显。对 sybil 攻击类型的识别有着比较高的准确率。但同时因为存在 Grid Sybil Attack 的攻击，使得一部分正常消息被误判为 Grid 类型消息。

多分类器参数和结果如表 6 所示：

Table 6. Multi-classifier parameters and results
表 6. 多分类器参数和结果

输入信号个数	隐藏层节点	激活函数	学习率	迭代次数
18	36	Relu	0.001	200
类型	准确率	召回率	F1 得分	支持数据数量
Genuine	0.878	0.850	0.864	2240
Grid	0.868	0.791	0.830	2342
DataReplay	0.853	0.751	0.802	2252
DosRandom	0.752	0.986	0.870	2278
DosDisruptive	0.724	0.653	0.689	2323

图 8 显示了使用 MLP 训练时各个特征的重要性，并以得分的形式给出。

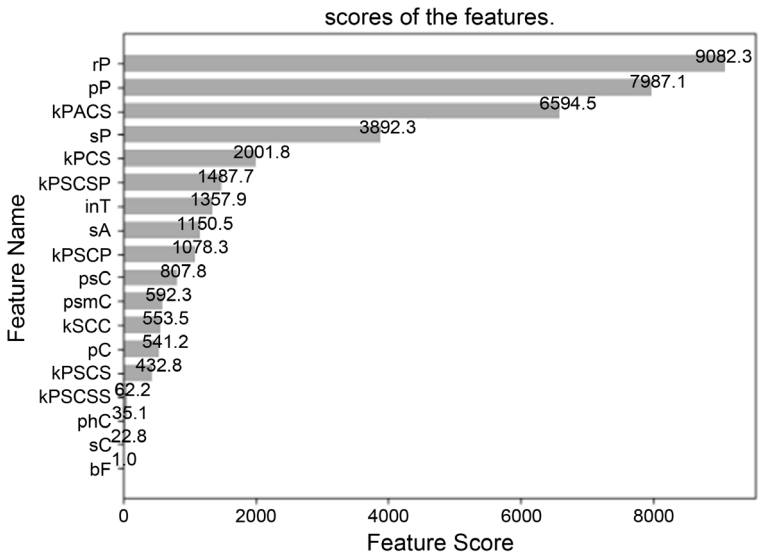

Figure 8. Two classifier feature score
图 8. 二分类器特征得分

由之前二分类器的结果分析得知，异常行为和正常消息具有比较大的区分度，使用 MLP 训练的结果能够识别异常行为。但在 sybil 的多分类识别中，通过对结果进行分析可以得到下述结论：真实 Data Replay 类型的消息容易被误判为 Dos Disruptive 类型，而 Dos Disruptive 类型的数据同样比较容易被误判为 Dos Random 类型和 Data Replay 类型。主要原因在于 Dos Disruptive Sybil Attack 是 Data Replay Sybil Attack 和 Dos Random Sybil Attack 的结合体，其特征的置信度均和前两种攻击有着比较大的联系。

5. 总结

实验结果表明，本文提出的基于流量统计的 MLP 入侵检测模型综合利用了范围性特征和合理性特征，在区分正常流量和 sybil 攻击上平均达到了 85%以上的准确率，同时验证了利用范围性特征和合理性特征检测 sybil 攻击行文的有效性，但在区分不同 sybil 攻击类型上，仍然需要进一步的改进，本文提出以下一些改进方向：1) 本次实验仅在模拟环境中进行训练和测试，今后有待在实际环境中做进一步的验证与完善；2) 在整体架构设计中部分模块解耦不充分、功能划分不够清晰，后续应对其进行完善；3) 本次实验过程中使用的神经网络较为简单，未在各个 sybil 攻击行为的分类上达到很好的效果，未来将尝试使用其他神经网络，或使用融合神经网络做进一步的测试。后续将主要围绕上述几个改进方向开展工作，期待有更进一步的研究成果。

致　谢

本论文是在我的导师马进老师的指导下完成的。导师渊博的专业知识，严肃的科学态度，严谨的治学精神，诲人不倦的高尚师德都对我产生了深远的影响。从课题的选择到项目的最终完成，马进老师都始终给予了我悉心的指导。此外，本论文的顺利完成，还离不开 F2MD 框架作者的帮助，在系统的完成过程中一直耐心地回答我的问题，再次向他表达诚挚的感谢。最后还需要感谢在本次论文过程中给予我支持和帮助的同学们，在奋斗的日子里有一群志同道合的人做伴实在是一件幸事。

参考文献

[1] Sakiz, F. and Sen, S. (2017) A Survey of Attacks and Detection Mechanisms on Intelligent Transportation Systems: VANETs and IoV. *Ad Hoc Networks*, **61**, 33-50. https://doi.org/10.1016/j.adhoc.2017.03.006

[2] Kamel, J., Haidar, F., Jemaa, I.B., Kaiser, A., Lonc, B. and Urien, P. (2019) A Misbehavior Authority System for Sybil Attack Detection in C-ITS. 2019 *IEEE* 10*th Annual Ubiquitous Computing*, *Electronics & Mobile Communication Conference* (*UEMCON*), New York, 10-12 October 2019, 1117-1123. https://doi.org/10.1109/UEMCON47517.2019.8993045

[3] Pouyan, A. and Alimohammadi, M. (2014) Sybil Attack Detection in Vehicular Networks. *Computer Science and Information Technology*, **2**, 197-202.

[4] Ghaleb, F.A., Zainal, A., Rassam, M.A. and Mohammed, F. (2017) An Effective Misbehavior Detection Model Using Artificial Neural Network for Vehicular Ad Hoc Network Applications. 2017 *IEEE Conference on Application*, *Information and Network Security* (*AINS*), Miri, 13-14 November 2017, 13-18. https://doi.org/10.1109/AINS.2017.8270417

[5] Kamel, J., Kaiser, A., Ben Jemaa, I., Cincilla, P. and Urien, P. (2019) CaTch: A Confidence Range Tolerant Misbehavior Detection Approach. 2019 *IEEE Wireless Communications and Networking Conference* (*WCNC*) (*IEEE WCNC* 2019), Marrakech, 15-18 April 2019, 1-8. https://doi.org/10.1109/WCNC.2019.8885740

[6] Kalman, R.E. (1960) A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, **82**, 35-45. https://doi.org/10.1115/1.3662552

[7] Framework for Misbehavior Detection (F2MD). (2019) F2MD Website. https://www.irt-systemx.fr/f2md