doi:10.3969/j.issn.1000-436x.2014.z2.032

车联网中基于神经网络的入侵检测方案

刘怡良[1]，石亚丽[1]，冯蒿[2]，王良民[1]
(1. 江苏大学 计算机科学与通信工程学院，江苏 镇江 212013；2. 国网三门峡供电公司，河南 三门峡 472000)

摘 要：车联网的入侵检测（IDS）可用于确认交通事件通知中描述的事件的真实性。当前车联网 IDS 多采用基于冗余数据的一致性检测方案，为降低 IDS 对冗余数据的依赖性，提出了一个基于神经网络的入侵检测方案。该方案可描述大量交通事件类型，并综合使用了反向传播（BP）和支持向量机（SVM）2 种学习算法。这 2 种算法分别适用于个人安全驾驶速度快与高效交通系统检测率高的应用。仿真实验和性能分析表明，本方案具有较快的入侵检测速度,且具有较高的检测率和较低的虚警率。

关键词：车联网；入侵检测；神经网络；BP；SVM
中图分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2014)Z2-0233-07

Intrusion detection scheme based on neural network in vehicle network

LIU Yi-liang[1],SHI Ya-li[1],FENG Hao[2],WANG Liang-min[1]
(1.School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China;
2.State Grid Power Company Sanmenxia, Sanmenxia 472000,China)

Abstract: Vehicle networking intrusion detection solutions (IDS) can be used to confirm the authenticity of the events described in the notice of traffic incidents. The current Vehicle networking IDS frequently use detection scheme based on the consistency of redundant data, to reduce dependence on redundant data, an intrusion detection scheme based on neural network is presented. The program can be described as a lot of traffic event types , and the integrated use of the back-propagation (BP) and support vector machine (SVM) two learning algorithms. The two algorithms respectively applicable to personal safety driving fast and efficient transportation system with high detection applications. Simulation results and performance analysis show that our scheme has a faster speed intrusion detection, and has a high detection rate and low false alarm rate.

Key words: vehicle networking; intrusion detection; neural network; BP; SVM

1 引言

车联网在提高道路安全、交通效率方面具有巨大潜能，在未来交通环境中将发挥越来越大的作用[1]。由于车联网与驾乘人员的生命安全切实相关，所以车联网安全极其重要。当前关于车联网安全的方案大多集中于建立公钥密码体制（PKI）实现实体认证[2~4]，可以有效地排除未授权的实体。但这些协议不能排除具有合法身份的内部攻击者，为此，有研究引入入侵检测系统（IDS）以及时发现合法用户的恶意行为来提升系统的安全性[5~8]。

当前车联网的 IDS 主要是以数据为中心的方案[5~8]，多数基于信息一致性机制。该类方案假设车辆可以收到描述同一事件的多个信息。车

收稿日期：2014-07-01
基金项目：国家自然科学基金资助项目（61272074）；江苏省自然科学基金资助项目（BK2011464）；江苏省青蓝工程优秀中青年学术带头人；镇江市工业支撑基金资助项目（GY2013030）
Foundation Items: The National Natural Science Foundation of China (61272074); The Natural Science Foundation of Jiangsu Province (BK2011464); Blue Project of Jiangsu Province Outstanding Young Academic Leaders; Zhenjiang City Industrial Support Project（GY2013030）

辆检测这些信息的一致性来发现虚假信息。然而，这种一致性机制依赖当前的路由协议以及节点数目。而在大部分场景下，没有充足的节点提供足够的冗余信息。

针对上述问题，提出了一种基于神经网络的入侵检测方案。该方案定义了一个自适应和自学习的模型。在训练阶段，实时地收集事件附近的有用模式，用于建立描述攻击行为的模式库。在入侵检测阶段，车辆收到信息后，通过共享的模式库学习发现不符合事实的模式来判定欺骗性信息。在收集事件信息时，为了降低对冗余数据的依赖性，使用了传感器来感知事件周围的有关真实信息。

本文的主要贡献如下。

1) 根据一般网络的入侵检测框架，定义了一个基于神经网络的 IDS 框架,设计了该框架下的入侵检测方案。

2) 设计了适用于个人安全驾驶应用的神经网络学习算法，用来向驾驶者提供碰撞警告或者刹车警告从而避免潜在的事故。

3) 设计了适用于高效交通系统应用的神经网络学习算法，通过共享交通信息来整体优化交通系统的效率。

2 相关工作

车联网的 IDS 是指通过收集和分析事件的信息，检测车联网中是否存在虚假信息和占用网络资源的攻击行为的技术。近期研究者们提出的方案[5~8],多数是基于数据一致性机制，另外也有一些基于车辆主体的判定方案。如信任机制[9]、投票机制[10]、统计机制[11]。

数据一致性机制源于 Golle 等[5]提出的理论框架,他们假设车联网中仅有少数车辆是恶意车辆。为了防止数据丢失，每个数据都通过不同的通道转发。有多个车辆观察到事件。在这 3 个前提下，他们提出了使用模型来确认事件的真实性。当多个通道传来的通知出现不一致的情况时，由一个预先建立的模型来解释这种不一致情况的原因。Petit 等[6]发现了原有的一致性机制无法提供实时性的入侵检测，他们通过邻居节点的实时警报，建立动态的模型。在这个动态的模型中，车辆可以通过阈值做出决策。Petit 等[7]提出了一个动态的阈值更新算法，为驾驶员提供了足够的通知反应时间。为了进一步分析现有的协议能否从数据冗余中得到数据一致性特征。Dietzel 等[8]针对现有的车联网路由协议做了大量的实验分析。结果显示现有的协议可以获得有效的数据冗余用来进行一致性检查。

其他机制有基于信任机制，Kim 等[9]提出了一个基于多尺度的信息分类模型，实现对不同种类信息的检测。分类模型基于 2 个原则：1) 事件的重要性；2) 对接收的信息的信心程度。车辆可以通过多个尺度来确认通知的真实性；基于投票机制，Raya 等[10]提出了一个精简的证书撤销列表（RC^2RL）用于排除恶意车辆。由于 RSU 覆盖范围有限，在正常车辆没有收到 RSU 广播的撤销信息前，他们提供了一个投票方案（LEAVE）暂时性地隔离异常节点，基于统计机制，Chun 等[11]提供了一个安全高效的概率统计协议。由多个车辆合作完成对事件的认证。每个车辆将结果反馈给 IDS 进行统计。由于车辆反馈的结果不一定真实，因此他们从概率的角度上分析了结果的可信度，并作为入侵判断依据。

总体来说，关于上述车联网 IDS，其主要目的是达到实时、通用、适用 3 个方面的性能。

1) 实时性：主要针对能否及时发现入侵行为并反馈信息的问题。在复杂的道路上，车辆需要 IDS 进行实时性的决策以处理道路上的突发事件。

2) 适用性：主要针对能否发现道路上未知异常和已知入侵的问题。未知异常识别基于一个正常的模型数据库，通过发现与正常模型的偏离定位异常；已知入侵检测通过预先获得的攻击特征检测入侵行为，主要是识别已知的攻击行为。

3) 通用性：主要针对网络适用性的问题。网络适用性支持信息单跳和多跳 2 种传播方式。通常，通知信息根据车辆数目和需要传播的距离分为单跳和多跳 2 种方式。在刹车警报等单跳应用中，车辆只需发送刹车警报通知自己附近的车辆；而在道路拥塞警报等多跳应用中，需要通过信息转发的方式将警报发送给多个网络跳数之外的车辆。

当前相关工作在上述三方面的性能如表 1 所示。

3 基于神经网络的 IDS 模型

3.1 系统模型

基于神经网络的 IDS 系统建立在如下车联网系统模型的基础之上。车联网的系统模型如图 1 所示。

表 1　　**各方案的性能**

机制	方案	实时性	已知入侵检测	未知异常检测	单跳	多跳
一致性	文献[5]	—	已知入侵检测	—	单跳	—
	文献[6~8]	实时性	已知入侵检测	未知异常检测	—	多跳
信任	文献[9]	实时性	已知入侵检测	—	单跳	—
投票	文献[10]	实时性	—	—	单跳	—
统计	文献[11]	实时性	已知入侵检测	未知异常检测	单跳	—

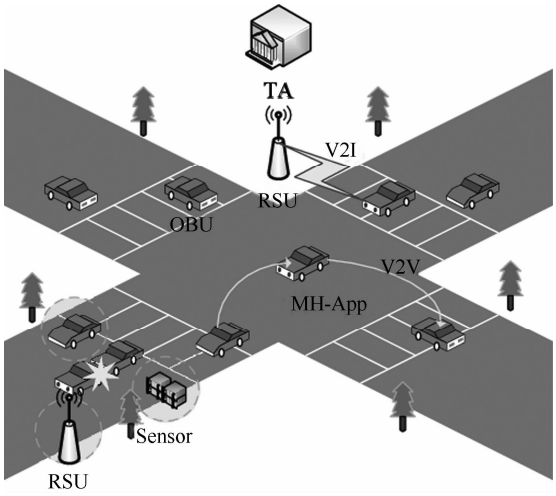


图 1　车联网的系统模型

由图 1 可知，该模型包含可信中心（TA）、路边基站（RUS）、车载设备（OBU）、传感器 4 个部分。模型中各部分的功能如下。

1) 可信中心（TA）：TA 用于向网络中的实体发布密钥信息。车辆发送的信息需要经过密钥的签署以进行车辆的身份认证，并保障信息的完整性。

2) 路边基站（RSU）：RSU 广范分布于网络中，可以与 TA 和 OBU 进行通信；同时 RSU 具有入侵检测功能以验证事件的真实性，并将结论发布给其他车辆。

3) 车载设备（OBU）：OBU 安装在车辆之上，通过 V2V 或 V2I 的方式进行通信。同时 OBU 也具有入侵检测功能，并周期性地广播道路相关信息。

4) 传感器：车联网系统模型中存在路边传感器和车载传感器 2 种传感器。路边传感器对路面或道路上的车辆的信息进行感知，当有车辆发生碰撞时，车辆周围的路边传感器会检测到路面的震动信号，周围车辆急剧减速的信号。车载传感器主要对车辆自身的状态及其行驶的路面进行感知，如道路结冰将会导致轮胎打滑，车辆在感知到路面信息后，及时改变驾驶行为，如放慢车速，不能急刹车等，防止因驾驶行为而导致车辆事故的发生。

3.2　攻击者模型

车联网中攻击者的行为主要包括 3 种形式[12]：数据篡改攻击、数据注入攻击、数据拘留攻击。

数据篡改攻击主要是篡改车联网中传播发送的信息。本文使用了签名算法对信息进行保护。

数据注入攻击主要体现在向车联网中注入虚假的或无效的通知，影响车联网的正常工作，甚至危害车联网用户的人身与财产安全。数据注入攻击存在 2 种非法注入的数据：1) 虚假数据，由攻击者随机生成的数据；2) 恶意赋值数据，由攻击者截获有效数据，直接赋值并注入网络。

数据拘留攻击主要是攻击者接收到安全数据分组后延误数据分组的转发，导致其他车辆没有足够的时间对紧急情况作出反应。

4　基于神经网络的 IDS 框架

在车联网系统基础之上设计了一个基于神经网络的 IDS 框架，图 2 为车联网中基于神经网络的 IDS 的框架总图。

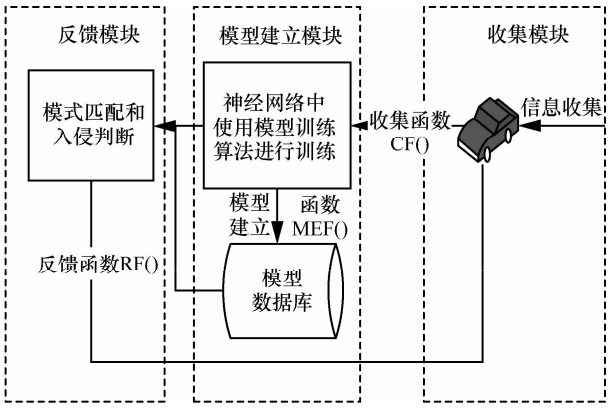


图 2　车联网中基于神经网络的 IDS 框架

由图 2 可知，该框架包含事件收集、数据库建立、反馈这 3 种功能模块。

1) 收集函数(CF)用于描述事件收集模块。

事件收集模块可以通过多种渠道收集事件的相关信息，然后将这些收集的信息传送至 IDS。收集函数 CF()表示如下

$$x_i = \mathrm{CF}(I_i), x_i \in I_i, i \in (0,1,2,\cdots,n) \quad (1)$$

式(1)中，n 表示事件 E_j 周围的信息数；I_i 表示事件 E_j 周围的第 i 个信息。收集函数 CF 选择信息中可以代表事件的特征项，如车辆的速度，加速度，位置等信息，将这些选择的特征项集合作为用于训练模型的输入集合 x。

2) 模型建立函数（MFF）用于描述数据库建立。

模块，数据库建立模块处理收集到的信息并建立车辆行为模型，同时储存这些模型用于后续的入侵检测。模型建立函数 MEF()表示如下

$$w_j = \mathrm{MEF}(x_1, x_2, x_3, \cdots, x_n), j \in (0,1,2,\cdots,m) \quad (2)$$

式(2)中，w_j 为事件 E_j 的权值。MEF 函数使用输入集合得到各个事件的权值，从而建立各个事件的模型数据库。模型数据库可以表示成以下数学模型

$$M = \begin{bmatrix} w_1, w_2, w_3, w_4, ..., w_m \\ E_1, E_2, E_3, E_4, ..., E_m \end{bmatrix} \quad (3)$$

3) 反馈函数（RF）用于描述反馈模块，反馈模块发送警报信息，排除被检测的攻击。它可以使车辆对事件的真实性做出决策。反馈函数 RF()表示如下

$$\mathrm{RF}(w_j, I_i') = E_j, i{+}{+} \quad (4)$$

式(4)中，I_i' 表示在入侵检测过程中，事件 E_j 周围的第 i 个信息。当 I_i' 与模型数据库中的模型进行匹配后，还需使用一个概率模型对匹配成功的传感信息数目进行统计。因为即使传感信息能成功匹配数据库中的某个模型，也不能确定该事件是真实的。

5 基于神经网络的 IDS 方案设计

在基于神经网络的 IDS 框架下设计了一个基于神经网络的 IDS 方案。该方案包括以下 3 个过程：信息收集、模型训练算法、模式匹配和入侵判断。

5.1 信息收集

为了确保收集的信息能够真实地反应事件，根据信息的产生时间和位置限定信息的接收范围，分别定义了一个时间域 TD 和空间域 SD 用于时间和位置的同步，假设 x_{i_1} 为第 i 个信息的时间戳，(x_{i_2}, x_{i_3}) 为其起始位置坐标，并将它们置于整个数据分组头部。收到信息的 IDS 会预先验证信息是否有效，再进行下一步的操作。

1) 时间域 TD 可以表示成以下数学模型

$$T - x_{i1} < \Delta t, \Delta t = \Delta t + D, \cdots, \Delta t + LD \quad (5)$$

其中，T 是指收到信息的时间，Δt 指预期的网络延迟。为了消除多跳过程中传输延迟产生的差异，采用同态时间戳机制，L 指跳步长度，D 指相邻节点间的信息延迟。

2) 设置（p_x, p_y）表示事件发生时的位置坐标。空间域 SD 可用表示成以下数学模型

$$\sqrt{(p_x - x_{i2})^2 + (p_y - x_{i3})^2} < \varepsilon \quad (6)$$

5.2 模型训练算法

车联网 IDS 的学习算法需要高效地泛化各种复杂的道路境况使入侵判断得出正确的结果，并且需要提供尽可能快的收敛性使未知的异常被快速泛化。为满足车联网的 IDS 需求，使用了 2 种著名的学习算法支持向量机（SVM）和反向传播算法（BP）。它们都具有极好的泛化性能，且能满足车联网中事件模型对及时性和自适应性的要求。

2 种学习算法用来训练特定的 w_k', β_k', b_k' 满足以下公式

$$\left\| H(w_1', \cdots, w_{n_h}', b_1', \cdots, b_{n_h}')\beta_k' - E \right\| = \min_{w_k, \beta_k, b_k} \left\| H(w_1, \cdots, w_{n_h}, b_1, \cdots, b_{n_h})\beta_k - E \right\| \quad (7)$$

式（7）中，w_k' 是连接第 k 个神经元和输入节点的权值向量，β_k' 是连接第 k 个神经元和输出节点的权值向量，b_k' 是第 k 个神经元的阈值。H 是一个优化算法，用来求权值和阈值的最优解。基于 BP 神经网络的学习算法使用梯度下降算法作为优化算法，基于 SVM 神经网络的学习算法使用拉格朗日乘子算法作为优化算法。

使用逼近 sin$C$ 函数的方式来评估 BP 和 SVM 的泛化性能，为了使回归更为逼真，在训练样本里加入了大量分布的噪音，但测试样本里无噪音。假设训练样本和测试样本数都为 5 000 个，它们通过随机抽取 x 轴的点获得。

BP 和 SVM 的对比如表 2 所示。

通过分析 BP 和 SVM 学习算法的性能，发现这 2 种学习算法适应于不同情况的应用。在跳步数较少的应用中，要求 IDS 可以快速反馈接收到的通知，

避免由于虚假通知造成事故或错误行驶。在表 2 中可清楚观察到 SVM 使用了 1 125.4 s 的 CPU 时间训练模型，而 BP 仅用了 19.63 s 的 CPU 时间就完成了相同的操作。BP 训练神经网络的速度比 SVM 大约快了 50 倍。因此，BP 学习算法更适用于个人安全驾驶的应用，用于向驾驶者提供碰撞警告或者刹车警告避免潜在的交通事故。在跳步数较多为远处车辆提供通知时，对反馈时间要求不高，而对检测率具有更高的要求。在表 2 中可从方差和均方差 2 个方面对比学习算法的泛化性能。其中方差表示结果与平均结果的偏移程度，均方差表示各个输出结果的偏离程度。对比显示 SVM 的方差与均方差都低于 BP，SVM 较 BP 具有更高的检测率。因此 SVM 更适合高效交通系统的应用，通过共享交通信息来优化交通。

表 2　　BP 和 SVM 的性能对比

学习算法		BP	SVM
训练	时间	19.63	1 125.4
	方差	0.118 3	0.113 7
	均方差	0.003 9	0.000 9
测试	时间	0.026	5.320
	方差	0.015 5	0.011 8
	均方差	0.005 6	0..001 7

根据上述分析，算法如下。

算法 1　Training algorithm

Begin procedure

1) For i=0:N,input(x_i, e_i)

2) if TD or SD is invalid then

3)　　Drop the message

4)　else if L>u(hop) then

5)　　Select the SVM to train the neural network

6)　　else

7)　　Select the BP to train the neural network

8)　　end if

9) end if

10) For i=0:m,out(w_i, e_i)

End procedure

5.3　模式匹配和入侵判断

车辆将收集到的事件周围的信息与训练得到的事件模型进行匹配，从而验证该事件通知的真实性。匹配过程可简洁地表示成以下数学模型

$$H(w,b,\beta,x_1',x_2',\cdots,x_s',e)=\begin{bmatrix} g(w\cdot x_1'+b)\beta-e \\ \cdots \\ g(w\cdot x_s'+b)\beta-e \end{bmatrix} \tag{8}$$

式（8）中，e 表示某个事件，x'表示入侵检测过程中收集的信息 I'通过收集函数 CF 得到的用于入侵检测的输入集。(w, β, b)表示该事件训练得到的权值。

理想状态下，如果匹配结果向量里所有的元素都等于 0，则该事件是真实。如果大部分元素都偏离了 0，则该事件是虚假的。

6　仿真实验和性能分析

6.1　完整性和检测时间

信息的完整性是指在传播过程中通知信息不会被攻击者篡改。为了保护信息的完整性，使用了基于身份的公钥密码体制。车辆用身份作为自己的私钥签署信息，信息的接收者使用车辆的公钥来验证信息，若公钥能解密打开信息，则这条信息是完整的。

一个完善的 IDS 的检测时间应包括信息认证开销和入侵检测开销 2 个部分。其中信息认证开销由签名方案的效率决定，使用 Wang 等[12]提出的批认证签名算法，算法认证 n 条信息的完整性共需要花费 2.9n+10 ms。入侵检测开销由训练算法的效率决定，图 3 给出了 IDS 的计算开销与收集的信息数目之间的关系。

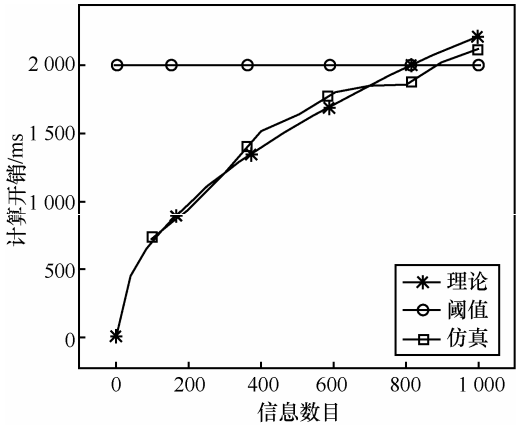
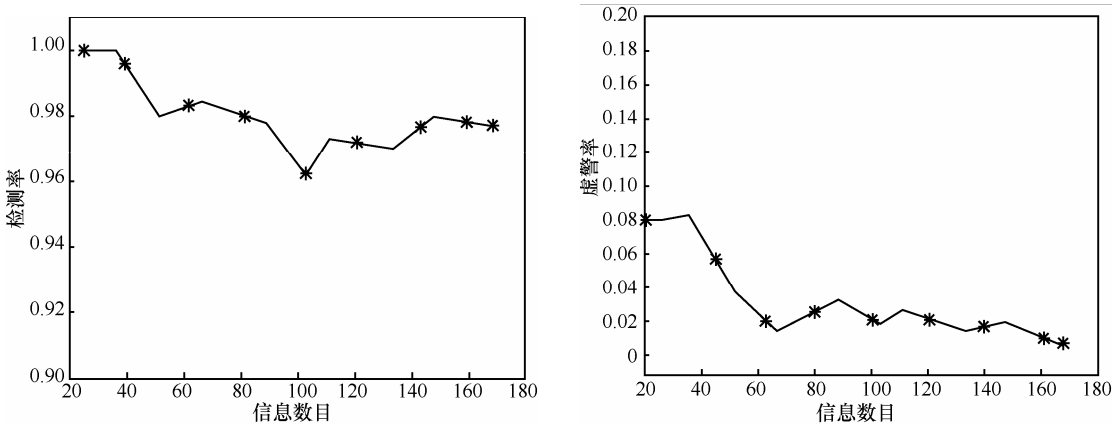


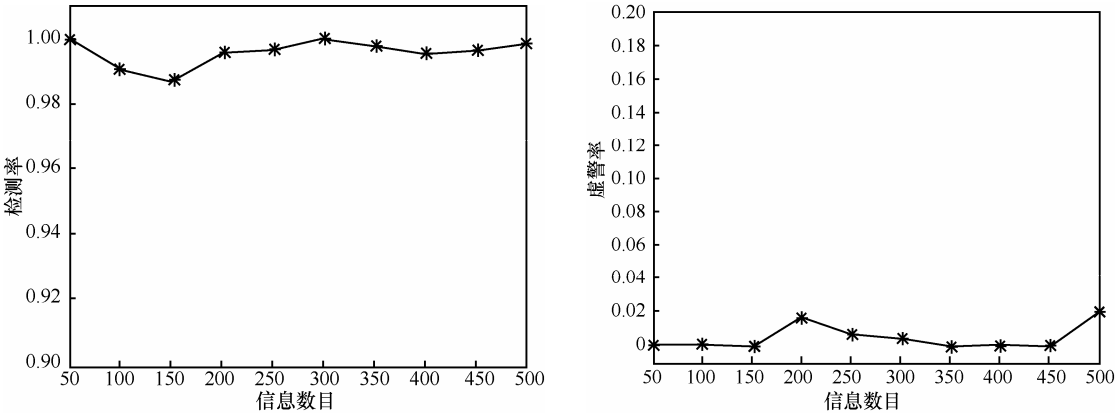
图 3　IDS 的计算开销与收集的信息数目之间的关系

从图 3 可以看出，随着信息数目增多，IDS 的理论计算开销和实际计算开销都明显的增加，且理论值与实际实验结果基本相符。根据 Wang 等[12]的结论，反馈时间上限应低于 2 000 ms，从图中可看出收集的信息数目低于 800 时基于神经网络的 IDS 具有较高的检测率。



(a) 基于 BP 神经网络的 IDS 的检测率　　(b) 基于 BP 神经网络的 IDS 的虚警率

图 4　基于 BP 神经网络的 IDS 的检测率与虚警率



(a) 基于 SVM 神经网络的 IDS 的检测率　　(b) 基于 SVM 神经网络的 IDS 的虚警率

图 5　基于 SVM 神经网络的 IDS 的检测率与虚警率

6.2　检测率和虚警率

关于检测结果的可信问题，通过检测 IDS 的检测率和虚警率来判断。检测率是指 IDS 系统能够正确检测出恶意车辆发出报警的概率，虚警率是指 IDS 系统在检测时错误检测出恶意车辆发出虚警的概率。

利用交通仿真软件 VanetMobisim 搭建道路仿真场景，模拟了道路上车辆节点的运动轨迹，并将生成的仿真数据作为基准数据集。随机抽取数据集中的部分数据作为训练集和检测集，并逐步增加虚假信息数据来检测方案的性能。图 4 和图 5 分别是基于 BP 神经网络的 IDS 和基于 SVM 神经网络的 IDS 的检测率和虚警率。

由图 4 和图 5 可知，本方案的 IDS 具有较高的检测率和较低的虚警率。

7　结束语

对车联网中的入侵检测进行了研究，为了不依赖冗余链路，依靠传感器得到事件周围的有关真实信息，并提出了基于神经网络的入侵检测方案。该方案使用了方向传播算法（BP）和支持向量机（SVM）2 种学习算法来建立入侵模型，这 2 种学习算法具有很好的泛化性能，能够满足车联网事件模型对及时性和自适应性的要求。通过仿真实验和数据分析得出 BP 训练神经网络的速度较快适用于个人安全驾驶的应用，SVM 的检测率较高只用于高效交通系统的应用。同时，性能分析表明本文方案具有较高的入侵检测效率，且具有较高的检测率和较低的虚警率。

参考文献：

[1] ISAAC J T, ZEADALLY S, CAMARA J S. Security attacks and solutions for vehicular ad hoc networks[J]. IET Communications, 2010, 4(7):894-903.

[2] RAYA M, HUBAUX J P. Securing vehicular ad hoc networks[J]. Journal of Computer Security, 2007, 15(1):39-68.

[3] STUDER A, BAI F, BELLUR B, *et al*. Flexible, extensible, and efficient VANET authentication[J]. Journal of Communications and Networks, 2009, 11(6):574-588.

[4] WASEF A, SHEN X M. EMAP: Expedite message authentication

protocol for vehicular ad hoc networks[J]. IEEE Transactions on Mobile Computing, 2013, 12(1):78-89.

[5] GOLLE P, GREENE D H, STADDON J. Detecting and correcting malicious data in VANETs[A]. Proceedings of the First International Workshop on Vehicular Ad Hoc Networks[C]. Philadelphia, PA, 2004. 29-37.

[6] PETIT J, FEIRI M, KARGL F. Spoofed data detection in VANETs using dynamic thresholds[A]. 2011 IEEE Vehicular Networking Conference (VNC)[C]. Amsterdam, 2011.25-32.

[7] PETIT J, MAMMERI Z. Dynamic consensus for secured vehicular ad hoc networks[A]. 2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications[C]. Wuhan, 2011.1-8.

[8] DIETZEL S, PETIT J, HEIJENK G, *et al*. Graph-based metrics for insider attack detection in VANET multihop data dissemination protocols[J]. IEEE Transactions on Vehicular Technology, 2013, 62(4): 1505-1518.

[9] KIM T H J, STUDER A, DUBEY R, *et al*. VANET alert endorsement using multi-source filters[A]. Proceeding of the Seventh International Workshop on Vehicular Ad Hoc Networks[C]. Chicago, IL, 2010. 51-60.

[10] RAYA M, PAPADIMITRATOS P, AAD I, *et al*. Eviction of misbehaving and faulty nodes in vehicular networks[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(8):1557-1568.

[11] HSIAO H C, STUDER A, DUBEY R, *et al*. Efficient and secure threshold-based event validation for VANETs[A]. Proceedings of the Fourth ACM Conference on Wireless Network Security[C].New York, NY, USA, 2011.163-174.

[12] LI X J, WANG L M. A rapid certification protocol from bilinear pairing for vehicular ad hoc networks[A]. Trust, Security and Privacy in Computing and Communcations[C]. 2012. 890-895.

作者简介：

刘怡良（1990-），男，江苏徐州人，江苏大学硕士生，主要研究方向为车联网、网络安全。

石亚丽（1992-），女，安徽芜湖人，江苏大学硕士生，主要研究方向为车联网安全。

冯嵩（1979-），男，河南新密人，国网三门峡供电公司工程师，主要研究方向为电力系统自动化、继电保护。

王良民（1977-），男，安徽潜山人，江苏大学教授、博士生导师，主要研究方向为物联网信息处理技术、物联网安全协议、车联网安全结构。

（上接第 232 页）

[7] OLIVIER D V, ANDERSON A, CORNEY M, *et al*. Multi-topic E-mail authorship attribution forensics[A]. ACM Conference on Computer Security Workshop on Data Mining for Security Applications, Philadelphia, 2001.

[8] PARIDHI J, PONNURANGAM K, ANUPAM J. @I seek 'fb.me': identifying users across multiple online social networks[A]. IW3C2[C]. 2013.

[9] ZI C, STEVEN G, HAINING W, *et al*. Who is tweeting on Twitter: human, bot, or cyborg?[A]. ACSAC '10: Proceedings of the 26th Annual Computer Security Applications Conference[C]. 2010.21-30.

作者简介：

卜俊丽（1990-），女，陕西大荔人，国防科学技术大学硕士生，主要研究方向为信息安全、数据挖掘与分析等。

彭灿（1982-），男，湖北武汉人，国防科技大学硕士生，主要研究方向为信息安全、人工智能。

郑毅（1989-），男，重庆人，国防科技大学硕士生，主要研究方向为数据挖掘。

黄九鸣（1981-），男，福建安溪人，博士，国防科学技术大学助理研究员，主要研究方向为文本挖掘、社交网络分析与大数据处理技术。

周斌（1971-），男，江西吉安人，博士，国防科学技术大学研究员，硕士生导师，主要研究方向为互联网数据分析与挖掘、信息检索等。