基于马尔可夫决策过程的入侵检测方法研究

董 凯,赵 旭
(西安工程大学 计算机科学学院,陕西 西安 710600)

摘 要:随着网络规模日益扩大,网络安全事件层出不穷,传统的网络入侵检测方法已不能满足当前网络的发展态势。为解决传统入侵检测方法中误报率过高、检测率和检测效率低等问题,提出了一种基于马尔可夫决策过程的入侵检测模型。在入侵检测系统内,根据马尔可夫的基本要素建立马尔可夫决策过程,采用模糊层次分析法为用户设置信用度,完成对用户信用度体系和数据库的构建,通过检测引擎学习得到马尔可夫决策过程的最优策略。在最优策略求解中采用策略迭代方法,其核心思想是给定当前策略函数进行状态价值函数 V 的评估,对状态价值函数采用贪心算法来提高策略函数,使得未来的回报最大同时输出最优价值函数。最后为了验证该方法的有效性,将 MDP-IDS 模型与支持向量机模型进行对比,实验结果表明 MDP-IDS 模型能够提高入侵检测率和检测效率,降低系统误报率。

关键词:马尔可夫决策过程;入侵检测;信用度;模糊层次分析法;贪心算法
中图分类号:TP309 **文献标识码**:A **文章编号**:1673-629X(2021)05-0131-06
doi:10.3969/j.issn.1673-629X.2021.05.023

Research on Intrusion Detection Method Based on Markov Decision Process

DONG Kai,ZHAO Xu
(School of Computer Science,Xi'an Polytechnic University,Xi'an 710600,China)

Abstract:With the increasing scale of the network and the emergence of network security events,traditional network intrusion detection methods can no longer meet the current development trend of the network. In order to solve the problems of high false alarm rate,low detection rate and low detection efficiency in traditional network intrusion detection methods,we propose an intrusion detection model based on Markov decision process. In the intrusion detection system,the Markov decision process is established according to the basic elements of Markov,the fuzzy analytic hierarchy process is used to set the credit for the user,thus completing the user credit system and database,and the optimal strategy of Markov decision process is obtained through the detection engine learning. The strategy iteration method is used in the optimal strategy solution. The core idea is to evaluate the state value function V given the current strategy function. The greedy algorithm is used to improve the strategy function for the state value function,so that the future return is maximized and the optimal value function is output. Finally,in order to verify the effectiveness of the proposed method,the MDP-IDS model is compared with the support vector machine model. The experiment shows that the MDP-IDS model can improve the intrusion detection rate and detection efficiency,and reduce the system false alarm rate.

Key words:Markov decision process;intrusion detection;credibility;fuzzy analytic hierarchy process;greedy algorithm

0 引 言

随着互联网的不断发展,互联网已经成为生活中不可缺少的一部分。网络技术在商业、经济、军事等各个领域都发挥着巨大作用。根据中国互联网络信息中心最新的统计报告显示,截至 2020 年 3 月,国内互联网普及率达 64.5%。与此同时带来的挑战也更加严峻,成千上万的网络攻击、网络安全事件层出不穷。如今在大数据时代背景下,网络给人们的生活提供了便利,但同时网络安全问题日益凸显,一旦出现安全事件,将造成不可挽回的经济损失、社会影响。所以事前主动检测、防御,对于网络稳定、可靠运行具有重要的意义。针对网络与信息安全的发展,目前主要通过防火墙技术、数据加密、访问控制、入侵检测等方法进一步提高网络的安全性。通过对比以上几种常见的措

收稿日期:2020-07-14 **修回日期**:2020-11-18
基金项目:陕西省科技计划项目(2019KRM153);西安市科技创新引导项目(201805030YD8CG14(8))
作者简介:董 凯(1996-),男,硕士研究生,研究方向为网络信息安全、机器学习;通信作者:赵 旭(1978-),男,副教授,博士,研究方向为网络安全、应急管理。

施,入侵检测方法具有相对较高的灵活性和拓展性。现如今的网络安全问题一般需要通过多种技术的组合来进行保护和监测,依靠传统的防御手段已经无法更好地处理千变万化的网络问题。在目前广泛应用的防御手段中,提前发现异常的入侵行为并及时处理,能更好地适应当前的网络常态。网络安全中能否检测出网络异常行为是至关重要的一个环节,网络异常行为检测作为防火墙的重要补充,在不影响网络性能的情况下完成了对网络安全性的分析,并实时阻止攻击行为破坏网络,保障网络运行的安全。因此,入侵检测技术在网络安全领域是不可或缺的一部分。

自 1980 年首次提出入侵检测模型至今,已有众多学者对入侵检测技术进行大量的探索与研究。例如,Ligun 等人[1]提出一种基于规则的方法,该方法以专家的经验作为规则编码成入侵检测系统的检验规则。但存在一个普遍的问题,需要设计一组能准确识别入侵行为的规则,然而如何设计一组合适的规则是一个尚未解决的问题。Lee 等人[2]提出了一种基于数据挖掘的方法,核心思想是从采集到的样本数据中获取高频的关键信息和联合规则,区别于人工设计的检验规则。该方法依靠大量的联合规则,使得系统过于复杂,检测效率低下。Siraj[3]提出了一种新的混合智能方法,通过在分类精度和处理时间上的改进实现入侵检测中的自动警报。Patra 等人[4]使用关联规则挖掘和多个最小支持等方法,应用于识别正常用户和异常用户。该方法将基于规则与基于数据挖掘[5]的方法结合起来,在一定程度上可以改进检测率,但是对于未知的入侵检测效果不太明显。传统的流量分析、特征提取等网络攻击流量检测方法和检测技术难以适应高速和大规模的互联网环境,无法高效、准确地检测攻击,必须对其加以改进。这些方法的研究与应用为该研究提供了有利的参考。

通过上述分析,该文提出了一种基于马尔可夫决策过程的入侵检测(intrusion detection based on Markov decision,MDP-IDS)模型。该模型结合马尔可夫的基本要素建立入侵检测的马尔可夫决策过程,通过检测引擎学习得到马尔可夫决策过程的最优策略进行决策。采用模糊层次分析法为用户设置信用度,当用户访问时,对于信用度高的用户直接放行,其他用户则采用马尔可夫决策过程进行判断。其判断过程是通过分析用户的历史行为信息、主机信息等来辨别用户信用度的大小,从而区分合法用户和恶意用户,保证合法用户的业务不受影响的同时阻断入侵主机。通过在网络节点建立分布式入侵检测服务器[6],尽早地将存在恶意入侵行为的用户进行拦截。各入侵检测服务器周期性地进行数据同步,从而达到数据一致。

1 MDP-IDS 检测模型

MDP-IDS 模型将马尔可夫决策应用于网络入侵检测,通过学习不断改进。其基本思想是:通过学习选择一个作用于环境的动作 a,环境接收该动作后状态会发生改变,接着会反馈一个奖励 r 给智能体,智能体根据强化信号和环境当前的状态再选择下一个动作 a,选择的原则是受奖赏概率值有没有逐渐增大,即如果智能体的某个决策行为导致来自外部的评价信号的增强,此后产生这个决策行为的趋势会逐渐地增强,反之系统产生这个动作的趋势便会逐渐地减弱。根据系统当前所处的环境来采取行动,以达到预期利益最大化的目的。其本质就是解决一个决策问题,即学会自动进行决策。

其中,智能体是进行状态感知、学习训练、动作选择的模块,环境是当前系统的状态,此处的环境是网络中的用户行为组成的共同体,状态是由一系列能描述环境的参数组成,动作是作用于环境进行状态的,奖励则是环境给予动作的奖励值。进而得到一个具有高效决策能力的入侵检测模型。其工作流程如图 1 所示。

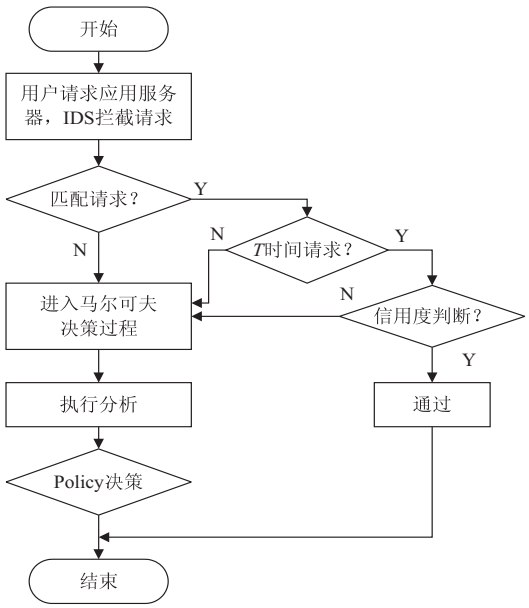

图 1 MDP-IDS 模型的工作流程

首先用户向应用服务器发送访问请求,入侵检测服务器拦截请求,检测引擎(detection engine)[7]查询数据库并对访问请求进行匹配,若查询不到请求记录或者记录不匹配,则将用户请求的信息记录到数据库;若信息存在,则不做任何操作。

若匹配到请求记录,则分两种情况处理。第一种情况是当用户在 T 时间周期内与服务器有过通信,则判断该用户的信用度是否满足条件。若信用度满足,则表示正常并且允许该用户请求服务器。若信用度不满足则拒绝该用户请求服务器。信用度较高表示该用户为诚实用户的可能性较高,反之则为恶意用户。第

二种情况是用户在 T 时间周期内未与服务器有过通信,则直接进入马尔可夫决策过程,Policy 函数根据分析的结果做出决策,并完成用户信用度的设置。

若无法匹配到请求,即该用户未与服务器建立过通信,亦采用马尔可夫决策过程进行决策。

1.1　马尔可夫决策过程建立

对于如何构建一个完整的马尔可夫决策过程,首先要在一个标准的马尔可夫决策过程中,将智能体设置为一个学习者,以便于获取外部环境的当前状态信息 s ,之后可以对环境采取试探行为 a ,并再次获取环境反馈对此动作的评价 r 和新的环境状态 s 。如果智能体在某动作 a 的作用下使得环境趋于正的奖励,那么智能体以后产生这个动作的趋势便会加强;反之,智能体产生这个动作的趋势将会减弱。且强化学习作为一种具有很强的决策能力的高阶机器学习方法,将其应用于入侵检测系统的构建当中,在学习系统的控制行为与环境反馈的状态及评价的反复的交互作用中,以学习的方式不断地修改从状态到动作的映射策略,即可以得到一个高效的、具有决策能力的入侵检测模型。

通过分析结合入侵检测与马尔可夫决策过程的特性,在入侵检测系统内建立学习环境。建立马尔可夫决策过程的 5 元组(S , A , P , R , γ),不断学习出最优策略[8],进而求出策略函数的最优解。建模的具体步骤如下:

(1)在入侵检测系统内设置入侵检测引擎为智能体(Agent)进行学习,即为动作的执行者;

(2)定义智能体的动作为 a ,动作空间 $A=\{a_1,a_2,a_3\cdots\}$ 是通过对 IP 数据库等入侵检测相关信息分析得到的动作集合。并且设置策略函数 π 为最优的检测方法;

(3)定义当前学习环境下的状态 s ,状态空间为 $S=\{\text{Normal},\text{Attack}\}$;

(4)定义该学习环境下的奖励函数 R 。智能体根据当前状态 s 给予动作进行奖励或者惩罚。作为入侵检测系统是否存在入侵,即奖励函数 R 为入侵系统的检测率[9];

(5)定义 γ 为折扣因子,通常采用折扣累计回报进行计算,且环境中的不确定性导致下一时刻的奖励权重小于当前时刻;

(6)模拟与环境相似的场景进行建模,进而学习到入侵检测的最优策略,然后利用递归的 Bellman 方程[10]进行求解。

1.2　累积回报和策略的表示

强化学习的最终目标是通过学习得到累积回报最大化的策略函数。而对于累积回报常用的是“ γ 折扣累计回报”方法。首先通过在环境中不断地尝试而得到一个策略函数[11](policy function),根据当前的策略函数,得到当前状态下要执行的动作。

采用随机性策略表示方法将策略函数表示为 π: $S\times A\rightarrow R$ 。

定义状态 s 下选择动作 a 的概率,即策略函数为: $\pi(a\mid s)=\mathbb{P}(A=a\mid S=s)$ 。且必须满足:

$$\sum_{a}\pi(s,a)=1 \tag{1}$$

同时还定义状态转移函数,即从当前状态中做出动作,使得转移到下一个状态,状态转移可以是确定的,也可以是随机的,状态转移的随机性是从入侵检测系统中来的。

$$p(s'\mid s,a)=\mathbb{P}[S'=s'\mid S=s,A=a] \tag{2}$$

累计奖励是指当前入侵检测环境,从 t 时刻开始的奖励全部加起来,而通常情况下采用折扣累计奖励,未来的不确定性使得 R_{t+1} 的权重低于 R_t 的权重。

定义折扣率 $r\in(0,1)$ 和累积奖励 U_t :

$$U_t=R_t+\gamma R_{t+1}+\gamma^2R_{t+2}+\gamma^3R_{t+3}+\cdots \tag{3}$$

累积奖励 U_t 存在两个随机性[11]:

(1)动作 a 的随机性,用状态 s 作为输入策略函数的输入,动作 a 是以概率分布的形式随机抽样得到的。

(2)下一个状态 s 的随机性,给定当前状态 s 和动作 a ,下一个状态 s 是随机的,状态转移函数 p 输出一个概率分布,环境从概率分布中抽样得到新的状态。对于任意时间 t ,奖励 R_t 取决于状态 S_t 和动作 A_t ,因此 U_t 的随机性是未来所有的状态和动作。

1.3　马尔可夫决策过程的求解

求解的核心思想就是找到一个最优策略函数,使得未来的回报最大,同时输出最优价值函数。以上过程即是马尔可夫的控制问题。在马尔可夫决策过程中,控制问题可以通过动态规划来求解,核心思想是把马尔可夫过程分解成每一个最佳的子结构。在 Bellman 方程中,包含两个函数:状态价值函数(state value function),表示状态上的累计奖励;动作价值函数(action value function),表示动作上的累计奖励。bellman 最优方程为:

$$V_*(s)=\max_a q(s,a)=\max_a\Big[R_s^a+\gamma\sum_{s\in S}P_{ss'}^aV_*(s')\Big] \tag{4}$$

$$q_*(s,a)=R_s^a+\gamma\sum_{s\in S}P_{ss'}^aV_*(s')=R_s^a+\gamma\sum_{s\in S}P_{ss'}^a\max_{a'}q_*(s',a') \tag{5}$$

$$\pi^*(a\mid s)=\begin{cases}1, & \text{if } a=\arg\max\limits_{a\in A}q_s(s,a)\\ 0, & \text{otherwise}\end{cases} \tag{6}$$

已知一个马尔可夫决策过程(S , A , P , R , γ),

在寻找最优策略的同时得到一个最佳的价值函数(optimal value function)。但在这种情况下,最优的价值函数是一致的,可能存在多个最优策略。而对于最优策略的收敛性,应满足以下条件 $\pi \geqslant \pi'$ if $V_n(s) \geqslant V_\pi(s), \forall s$。

进而通过对 q^* 求最大化得到最优策略:求解方法见公式(6)。

该文采用策略迭代的方法进行策略求解。求解过程可以分为两步:

(1)策略评估:给定当前策略函数然后计算状态价值函数 V;

(2)策略提升:对状态价值函数 V 采用贪心算法[12]来提高策略函数。

通过以上步骤,完成了 MDP-IDS 的模型构建。策略迭代过程如图 2 所示。

$$q^{\pi_i}(s,a) = R(s,a) + \gamma \sum_{s' \in S} P(s' \mid s,a) v^{\pi_i}(s') \quad (7)$$

$$\pi_{i+1}(s) = \arg\max_a q^{\pi_i}(s,a) \quad (8)$$

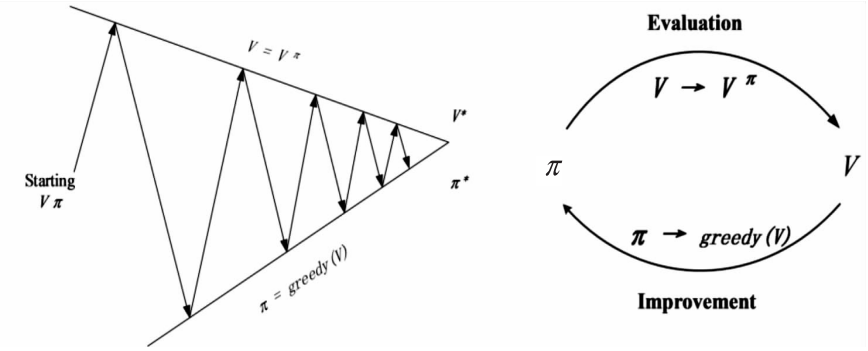

图 2 策略迭代过程示意图

1.4 信用度体系构建

用户信用度是对用户可靠性的衡量指标。信用度较高表示该用户为诚实用户较高,反之则为恶意用户。在信用度体系中使用模糊层次分析法(fuzzy analytic hierarchy process,FAHP)[13]对用户信用度进行评估,模糊层次分析法及计算过程层次分析法(AHP)是一种定性与定量相结合的多目标决策方法,能够有效分析目标准则体系层次间的非序列关系,有效地综合测度决策者的判断和比较。

模糊层次分析法的基本思想是根据多目标评价问题的性质和总目标,把问题本身按层次进行分解,构成一个由下而上的梯阶层次结构。因此在运用 AHP 决策时,大体上可以可分为以下四个步骤:问题分析,确定系统中各因素之间的因果关系,对决策问题的各种要素建立多层次递阶结构模型;对同一等级的要素以上一级的要素为准则进行两两比较,并根据评定尺度确定其相对重要程度,最后据此建立模糊判断矩阵;通过一定计算,确定各要素的相对重要度;通过综合重要度的计算,对所有的替代方案进行优先排序,从而为决策人选择最优方案提供科学的决策依据。再由 Policy 函数输出当前用户的信用度。首先将用户行为进行细分,每个行为即是一个特性,再将特性进行分类,将用户行为信用评估转换成为信用加权问题。分为以下几个步骤:

(1)从构建的数据库中获得数据并且是初始数据,为了便于数值计算和用户信用度评估,将数据全部规范化,表示为矩阵 $\boldsymbol{E} = (e_{ij})_{mn}$。为了获得初始判断矩阵 $\boldsymbol{EQ} = (eq_{ij})_{m\times m}$,有 m 个矩阵 $\boldsymbol{W} = (w_1, w_2, \cdots, w_m)^{\mathrm{T}}$,将矩阵集中在 e_i 和 e_j 进行对比:

$$eq_{ij} = \begin{cases} 0, & e_i < e_j \\ 0.5, & e_i = e_j \\ 1, & e_i > e_j \end{cases} \quad (9)$$

(2)将初始化判断矩阵转化为模糊一致的矩阵 $\boldsymbol{Q} = (q_{ij})_{m\times m}$,其中 $q_{ij} = \dfrac{q_i - q_j}{2m} + 0.5$, $q_i = \sum_{k=1}^{m} eq_{ik}$;

(3)计算某个特性的权重向量 $\boldsymbol{W} = (w_1, w_2, \cdots, w_m)^{\mathrm{T}}$,其中 $w_i = \dfrac{\sum_{k=1}^{m} q_{ik} - 0.5}{m(m-1)/2}$;

(4)计算用户行为特征的评估值矩阵,可根据 $E \times W^{\mathrm{T}}$ 得到特征值评估矩阵 $\boldsymbol{F} = (f_1, f_2, \cdots, f_n)$;即可得到当前用户的信用度为:

$$\text{Credit\_Value} = 1 - F \times \boldsymbol{W}_f^{\mathrm{T}} = 1 - \sum_{i=1}^{n} f_i w_i \quad (10)$$

1.5 数据库的构建

此模型在入侵检测服务器按照一定的概率 f 统计成功建立访问连接的 IP 数据包的源地址,包含 TCP、UDP、ICMP 数据包,将数据包的源地址 IP_Client_Source 存入表中。

在数据库模块中,按照以下的步骤建表。根据访问用户的历史访问记录,为每个目的服务器建立一张表。在每个服务器的表中,依据源地址进行聚集[14]以完成不同用户区分,完成用户信息的录入。其中数据库表的各个字段见表 1。

表 1　数据库表字段

字段	描述
IP_Server_Destination	此 IP 数据包的目的服务器地址
IP_Client_Source	此 IP 数据包的源地址
Access_Time_First	此 IP 数据包第一次记录的时间戳
Access_Time_Last	此 IP 数据包上一次记录的时间戳
Count_Mark	此 IP 数据包被记录的次数
Credit_Value	此 IP 数据包的信用度
Historical_Attack _Status	此 IP 数据包是否发生过入侵
TCP_Amount	TCP 数据包的数目
UDP_Amount	UDP 数据包的数目
ICMP_Amount	ICMP 数据包的数目

MDP-IDS 模型要求实时进行数据库的更新,及时地把请求添加到数据库中,以及对请求的相关字段数据的更新。为了尽早地将存在恶意入侵行为的用户进行拦截,同时减小服务器的压力加快检测效率,在网络中建立多个分布式入侵检测服务器,各分布式服务器通过 Raft[15] 算法进行同步,进而保证数据的一致性。

2　仿真分析

该文在 KDD CUP99[16] 基础上利用 Matlab 2018a 进行了仿真实验,通过与基于支持向量机(support vector machine,SVM)[16] 入侵检测方法进行对比,验证该方法的有效性。KDD99 数据集共 500 余万条,提供了 10% 的用于训练的子集和测试的子集。首先采用 one-hot[17] 方法对 10% 的训练集数据进行预处理,对数据的预处理结果会影响入侵检测实验的效果。利用 Python3 对训练数据集进行预处理,即进行字符型特征与数值型的转化。

为了对文中所提方法进行衡量,定义 TP(true positive)、FP(false positive)、FN(false negative)、TN(true negative),其中 DR(detection rate)、FAR(false alarm rate)、DT(detection time)用于结果评价,DR 是检测出的已知攻击数量与总数量的比率,FAR 是误判别攻击数量与正常数量的比例[18],DT 是检测引擎处理任务所需要的时间。DR 和 FAR 用以下公式来表示:

$$DR = \frac{TP}{TP + TN} \tag{11}$$

$$FAR = \frac{FP}{TN + FP} \tag{12}$$

检测引擎分别对 TCP、UDP、ICMP 进行实验分析,并与支持向量机的入侵检测方法进行对比。对比实验结果如表 2 所示。

表 2　DR 数据对比　　%

	MDP-IDS	SVM
ICMP	98.87	98.16
TCP	99.35	97.85
UDP	98.61	97.76

表 3　FAR 数据对比　　%

	MDP-IDS	SVM
ICMP	0.56	0.69
TCP	0.41	0.46
UDP	0.37	0.33

表 4　DT 数据对比　　S

	MDP-IDS	SVM
ICMP	9.35	10.46
TCP	7.4	9.73
UDP	6.67	7.65

将该文提出的 MDP-IDS 模型与支持向量机的入侵检测方法相比较,如上表所示,MDP-IDS 方法在检测率有明显的优势,平均检测率提高 1.02%,平均误报率下降 0.08%,系统检测时间效率提高 15.8%。

3　结束语

该文研究了基于马尔可夫决策过程的入侵检测方法,为网络入侵检测提供了一种新的思路,建立了 MDP-IDS 模型。通过检测引擎分析用户信用度、行为等信息,利用马尔可夫模型进行异常入侵行为自动决策,从而更好地区分合法用户和恶意用户。将用户信用度引入到入侵检测系统中,使用模糊层次分析法对用户信用度进行设置,使得用户信用度计算更加合理。实验结果证明 MDP-IDS 模型能够缩短入侵检测时间,提高系统的整体性能。网络异常行为检测的结果通常将网络行为分类两大类:正常和异常,异常行为又

可以分成很多小类，如 DOS、Probe、U2R、R2L 等常见攻击类型，因此下一步工作可以针对网络异常行为检测的多分类问题进行深入研究。

参考文献：

[1] LIGUN K，KEMMERER R. State transition analysis：a rule-based intrusion detection approach[J]. IEEE Transactions on Software Engineering，1995，21(3)：181-199.

[2] LEE W，STOLFO S J，MOK K W. A data mining framework for building intrusion detection models[C]//Proceedings of the 1999 IEEE symposium on security and privacy (Cat. No. 99CB36344). Oakland，CA，USA：IEEE，1999：20-32.

[3] SIRAJ M M，MAAROF A，ZAITON S. A hybrid intelligent approach for automated alert clustering and filtering in intrusion alert analysis[J]. International Journal of Computer Theory and Engineering，2009，1(5)：539-545.

[4] PANDA M，PATRA M R. Mining association rules for constructing a network intrusion detection model[J]. International Journal of Applied Engineering Research，2009，4(3)：381-398.

[5] 周立军，张 杰，吕海燕. 基于数据挖掘技术的网络入侵检测技术研究[J]. 现代电子技术，2016，39(6)：10-13.

[6] 孙 乔，邓卜侨，王志强，等. 一种基于分布式服务器集群的可扩展负载均衡策略技术[J]. 电信科学，2017，33(9)：190-196.

[7] 尚文利，安攀峰，万 明，等. 工业控制系统入侵检测技术的研究及发展综述[J]. 计算机应用研究，2017，34(2)：328-333.

[8] ZHAO Zongya，WANG Chang，YUAN Qingli，et al. Dynamic changes of brain networks during feedback-related processing of reinforcement learning in schizophrenia[J]. Brain Research，2020，1746：146979.

[9] 毕 猛，王安迪，徐 剑，等. 基于离散马尔可夫链的数据库用户异常行为检测[J]. 沈阳工业大学学报，2018，40(1)：70-76.

[10] 陈映瞳. Bellman 不等式的证明、应用及推广[J]. 高等数学研究，2009，12(3)：56-59.

[11] 刘 全，翟建伟，章宗长，等. 深度强化学习综述[J]. 计算机学报，2018，41(1)：1-27.

[12] 邹哲讷. 贪心算法及其应用[J]. 计算机光盘软件与应用，2015(3)：85-86.

[13] 张 凯，潘晓中. 云计算下基于用户行为信任的访问控制模型[J]. 计算机应用，2014，34(4)：1051-1054.

[14] 王永利，徐宏炳，董逸生，等. 分布式数据流增量聚集[J]. 计算机研究与发展，2006，43(3)：509-515.

[15] 陈 陆，黄树成，徐克辉. 改进的 Raft 一致性算法及其研究[J]. 江苏科技大学学报：自然科学版，2018，32(4)：559-563.

[16] LI Y，XIA J，ZHANG S，et al. An efficient intrusion detection system based on support vector machines and gradually feature removal method[J]. Expert Systems with Applications，2012，39(1)：424-430.

[17] 傅依娴，芦天亮，马泽良. 基于 One-Hot 的 CNN 恶意代码检测技术[J]. 计算机应用与软件，2020，37(1)：304-308.

[18] 王洁松，张小飞. KDDCUP99 网络入侵检测数据的分析和预处理[J]. 科技信息，2008(15)：79-80.

(上接第 130 页)

niques：a study[J]. Digital Signal Processing，2010，20(6)：1758-1770.

[9] JUNG K H. A survey of reversible data hiding methods in dual images[J]. IETE Technical Review，2016，33(4)：1-12.

[10] CHANG C C，LIN M H，HU Y C. A fast and secure image hiding scheme based on LSB substitution[J]. International Journal of Pattern Recognition and Artificial Intelligence，2002，16(4)：399-416.

[11] 王继军. 利用差值扩展和直方图平移的可逆数字水印算法[J]. 小型微型计算机系统，2014，35(5)：1192-1195.

[12] HUANG F J，QU X C，KIM H J，et al. Reversible data hiding in JPEG images[J]. IEEE Transactions on Circuits and Systems for Video Technology，2016，26(9)：1610-1621.

[13] WEDAJ F T，KIM S K，HUANG H J，et al. Improved reversible data hiding in JPEG images based on new coefficient selection strategy[J]. EURASIP Journal on Image and Video Processing，2017，63(1)：1-11.

[14] WENG S W，ZHANG G H，PAN J S H，et al. Optimal PPVO-based reversible data hiding[J]. Journal of Visual Communication and Image Representation，2017，48：317-328.

[15] LI X L，LI J，LI B，et al. High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion[J]. Signal Processing，2013，93(1)：198-205.

[16] 李桂芸，邓桂英，赵逢禹. 一种基于 LSB 图像信息隐藏的改进算法[J]. 计算机系统应用，2012，21(4)：156-160.

[17] 任克强，肖璐瑶. 融合 CFT 和 LSB 的高容量可逆数据隐藏[J]. 液晶与显示，2019，34(4)：410-416.