基于无监督多源数据特征解析的网络威胁态势评估

杨宏宇，王峰岩
（中国民航大学计算机科学与技术学院，天津 300300）

摘 要：针对监督式神经网络测试网络威胁时需根据数据类别标记进行建模的局限性，提出了一种基于无监督多源数据特征解析的网络威胁态势评估方法。首先，设计了一个面向安全威胁评估的变分自动编码器−生成式对抗网络（V-G），将只包含正常网络流量的训练数据集输入 V-G 的网络集合层进行模型训练，并计算各层网络输出的重构误差。然后，通过输出层的三层变分自动编码器重构误差学习并获取训练异常阈值，使用包含异常网络流量的测试数据集测试分组威胁并统计每组测试的威胁发生概率。最后，根据威胁发生概率确定网络安全威胁严重度，结合威胁影响度计算威胁态势值以获取网络威胁态势。仿真实验结果表明，所提方法对网络威胁具有较强的表征能力，能够有效直观地评估网络威胁的整体态势。

关键词：无监督；多源数据特征解析；变分自动编码器−生成式对抗网络；威胁发生概率；威胁态势评估
中图分类号：TP309
文献标识码：A
doi: 10.11959/j.issn.1000−436x.2020015

Network threat situation assessment based on unsupervised multi-source data feature analysis

YANG Hongyu, WANG Fengyan
School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China

Abstract: Aiming at the limitations of supervised neural network in the network threat testing task relying on data category tagging, a network threat situation evaluation method based on unsupervised multi-source data feature analysis was proposed. Firstly, a variant auto encoder-generative adversarial network (V-G) for security threat assessment was designed. The training data set containing only normal network traffic was input to the network collection layer of V-G to perform the model training, and the reconstruction error of the network output of each layer was calculated. Then, the reconstruction error learning was performed by the three-layer variation automatic encoder of the output layer, and the training abnormal threshold was obtained. The packet threat was tested by using the test data set containing the abnormal network traffic, and the probability of occurrence of the threat of each group of tests was counted. Finally, the severity of the network security threat was determined according to the probability of threat occurrence, and the threat situation value was calculated according to the threat impact to obtain the network threat situation. The simulation results show that the proposed method has strong characterization ability for network threats, and can effectively and intuitively evaluate the overall situation of network threat.

Key words: unsupervised, multi-source data feature analysis, V-G, threat probability, threat situation assessment

收稿日期：2019–10–17；**修回日期**：2019–12–17
基金项目：国家自然科学基金民航联合研究基金资助项目（No.U1833107）
Foundation Item: The Civil Aviation Joint Research Fund Project of National Natural Science Foundation of China (No.U1833107)

1 引言

网络技术和信息技术的快速发展与更替，使网络攻击特征多元化问题日益突出。面对特征多元化的网络攻击威胁，构建动态的主动网络防御体系成为实现网络安全保障的重要手段，而高效、直观的网络威胁态势评估方法则是网络动态主动防御体系架构的基础。目前，针对延展性强、稳定高效且高稳健性的网络威胁态势评估方法研究已成为重要课题，在威胁态势评估领域，基于数学模型、基于逻辑规则推理和基于深度学习技术 3 类方法应用较广泛。

基于数学模型的评估方法通过对网络威胁因素的综合分析，建立威胁指标集与威胁态势的对应关系，并将威胁态势评估问题归属于多指标综合评价或者多属性集合等问题。Yang 等[1]提出一种云计算风险评估模型，利用马尔可夫链（MC, Markov chain）模型描述随机风险环境并通过信息熵（IE, information entropy）度量风险值大小。Wang 等[2]将层次分析法（AHP, analytic hierarchy process）与情境评估的层次模型相结合，利用 D-S 证据理论融合多源设备的模糊结果，解决仅依靠单一信息源进行威胁评估时准确度偏差大的问题。由于基于数学模型的评估方法受主观因素影响较大，且没有客观统一的标准定义变量，因此通常无法取得较理想的评估效果。

基于逻辑规则推理的评估方法利用先验知识的统计特性，结合专家知识和经验数据库搭建模型，采用逻辑规则推理方式对网络威胁态势进行评估。Sallam[3]通过基于最大−最小模糊推理（FR, fuzzy reasoning）引擎的模糊逻辑技术识别网络潜在威胁，根据攻击者的总体能力、攻击成功的总体可能性和攻击影响这 3 个子模糊推理系统评估网络安全风险。文志诚等[4]通过分级朴素贝叶斯分类器融合信息源，使用数理统计方法融合各安全评估指数，对网络安全态势进行量化评估。随着网络场景的变化和网络信息的多元化，在网络安全态势评估时，基于逻辑规则推理的评估方法的局限性和不足愈发突出，尤其在面对新型网络威胁时不能及时做出反馈，无法满足任务处理需求，导致评估效率降低。

由于具有高效、易实现的特点，基于深度学习的评估方法近年来得到广泛应用。该类方法利用训练数据集对特定神经网络模型进行训练，并通过对网络参数的不断优化使模型收敛达到最优，然后根据分类和测试结果进行网络威胁态势评估。Feng 等[5]从原始时间序列网络数据中提取内部和外部信息特征，并将提取的特征在深度递归神经网络（RNN, recursive neural network）模型进行训练和验证，具有较高的预测准确性和稳健性。He 等[6]将小波神经网络（WNN, wavelet neural network）与最大重叠离散小波变换（MODWT, maximal overlap discrete wavelet transform）方法相结合，通过数据驱动方法提出网络安全态势预测模型。当面对海量网络安全数据时，基于深度学习的评估方法由于缺乏足够的先验知识和既定的数据类别标注准则，人工类别标注的任务量大且成本高，因此基于数据标签的监督式数据建模方式已逐渐无法适用于特定网络场景。

针对上述方法的不足和多源数据环境下的网络威胁态势问题，无监督学习（UL, unsupervised learning）提供了一种可行的解决思路。其主要特点是不需要人为标注数据类别，而是直接对预处理后的数据进行特征学习和建模。为实现多源数据环境下的网络威胁态势精准有效评估，本文提出一种基于无监督多源数据特征解析的网络威胁态势评估方法。从多源数据中选取特征并生成训练集，采用变分自动编码器−生成式对抗网络（V-G, variational auto encoder-generative adversarial networks）对训练集数据进行聚类解析，通过三层变分自动编码器计算训练误差阈值，利用异常流量数据集进行威胁测试，并根据威胁态势值计算结果对网络安全态势进行量化评估。实验结果表明，本文方法对网络威胁具有较好的评估效果，并且在面对网络威胁时具有较强的网络威胁表征能力，可在不依靠数据标签的情况下对网络威胁态势进行有效评估。

2 典型的无监督生成网络模型

2.1 变分自动编码器和生成式对抗网络

自动编码器（AE, auto-encoder）和变分自动编码器（VAE, variational auto-encoder）[7]均由编码器（encoder）和解码器（decoder）构成。两者的主要区别是，VAE 对编码器添加了“噪声”约束，迫使编码器产生服从单位高斯分布的潜在变量（LV, latent variable）集合。

AE 的网络结构如图 1 所示，VAE 的网络结构

如图 2 所示。

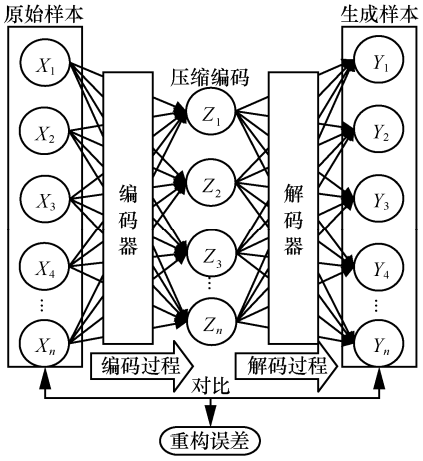


图 1　AE 的网络结构

由图 1 和图 2 对比可知，VAE 在编码过程中，要求原始样本集合 $X=\{X_1,X_2,X_3,\cdots,X_n\}$中的每一个样本 X_k均服从正态分布 $N(\mu, \sigma^2)$，即通过内部神经网络对任一样本 X_k的均值 μ 和方差 σ^2进行拟合，然后从所得正态分布中采样获取一个潜在变量集合 $Z=\{Z_1,Z_2,Z_3,\cdots,Z_n\}$，其中元素 Z_k服从多元标准正态分布 $N(0, I)$。在解码过程中，Z 通过解码器生成样本集合 $Y=\{Y_1,Y_2,Y_3,\cdots,Y_n\}$，然后使用距离函数统计生成样本集合 Y 和原始样本集合 X 的相似度大小，通过计算可获得整体数据元素的重构误差 loss。

生成式对抗网络（GAN, generative adversarial network）[8] 由生成器（generator）和判别器（discriminator）构成，它是无监督学习领域最具发展前景的深度生成网络模型之一。GAN 的网络结构如图 3 所示。

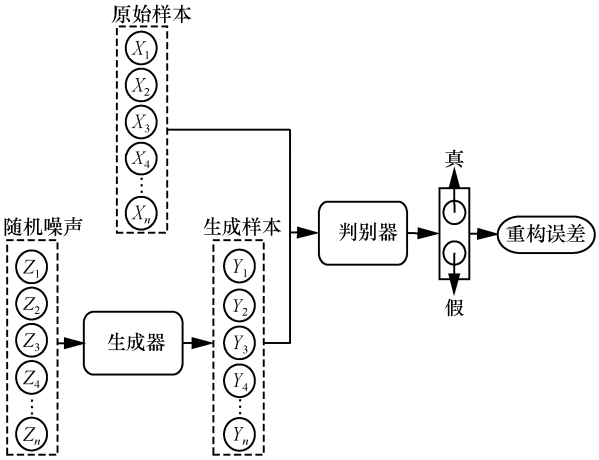


图 3　GAN 的网络结构

如图 3 所示，GAN 训练过程中，生成器首先学习通过先验分布直接采样获取的随机噪声集合 $Z=\{Z_1,Z_2,Z_3,\cdots,Z_n\}$的概率分布特性，然后尽量生成与原始样本集合 $X=\{X_1,X_2,X_3,\cdots,X_n\}$“完全”相似的数据集 $Y=\{Y_1,Y_2,Y_3,\cdots,Y_n\}$，以“欺骗”判别器；判别器则负责判别生成器所生成的数据集合 Y 与原始样本集合 X 的相似程度。判别器每进行一次相似度判别，均会输出[0,1]内的一个标量，标量越趋近于 0，生成样本 Y_k被判别为真实数据的概率越小；标量越趋近于 1，生成样本 Y_k被判别为真实数据的概

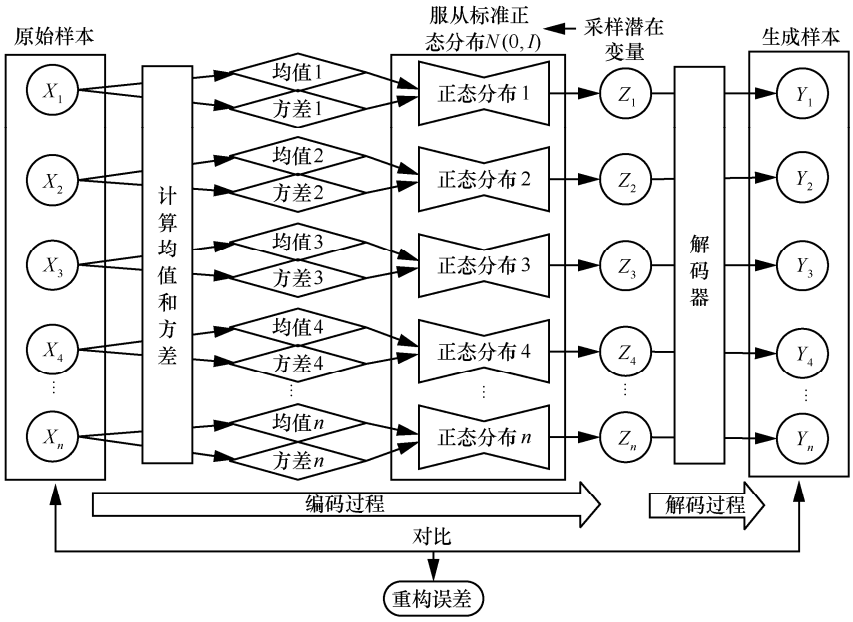


图 2　VAE 的网络结构

率越大。生成器和判别器由此构成一个动态的“博弈过程”，经过反复博弈，生成器逐渐学习到数据的分布特性，当达到纳什平衡点（NASH=0.5）时，生成器能够通过随机噪声集合 Z 准确还原出与原始样本集合 X 相似度极高的样本集合 Y，此时判别器无法区分真实数据与生成数据的真假，训练结束。

2.2 V-G 模型设计

通过对 2.1 节中 2 种典型的无监督网络模型的研究，对 VAE 和 GAN 这 2 种无监督模型的优缺点进行分析，结果如表 1 所示。

由表 1 可知，VAE 在编码过程中能够学习数据的先验分布，样本生成多样性性能良好，但在度量生成样本与原始样本的相似度时，只能使用均方根误差（MSE，mean squared error）等函数对数据元素间的相似度误差进行粗略计算，而无法采取更合理的相似度度量策略，这降低了样本匹配的准确度。GAN 在通过判别器进行样本相似度判别时，对生成样本和原始样本的判别标准很高，但 GAN 的生成器在生成样本时，由于其本身没有添加任何条件约束，因此解空间极大，对真实样本分布的拟合不易收敛到一个较好的结果。此外，由于 GAN 在生成样本过程中容易出现输入的多个随机噪声样本对应同一类生成样本的情况，因此容易出现生成样本的多样性减少和陷入模式坍塌的情形。为进行优势互补，本文将 VAE 的编码器与 GAN 的判别器进行结合，提出一个 V-G 模型，该模型对原始样本的映射能力较强，能够保证样本生成的多样性不受限制，同时满足高精确度的样本相似度判别要求，V-G 的网络结构如图 4 所示。

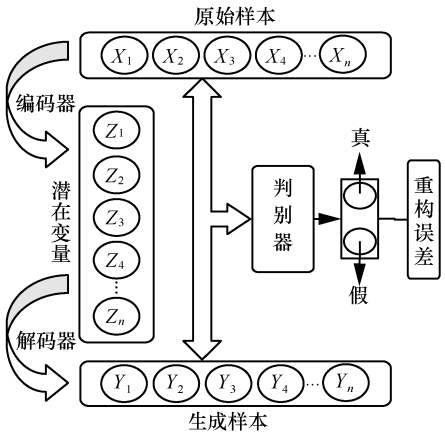


图 4 V-G 的网络结构

由图 4 可见，VAE 的解码器与 GAN 的生成器由于共享参数，可视为同一部分。将其应用到网络安全数据集中，则 GAN 判别器中的特征学习结果将作为 VAE 重构样本数据的输入项。此外，在相似度度量方面，由原本 VAE 的元素误差度量转变为 GAN 的特征误差度量，更易获取数据的分布特性。

本文所提 V-G 模型主要用于网络威胁测试，其应用对象主要为基于主机、网络、服务器等终端产生的多源异构网络流量数据。由于 V-G 模型独特的结构优势，其在进行模型训练时能够有效提取数据特征信息，提升聚类准确率，因此能够取得较高的威胁测试准确率。

3 网络威胁态势评估架构

3.1 架构设计

为了快速发现并及时处理复杂多变的网络攻击特征，减少人工对网络安全信息数据的干预度，及时发现网络威胁和安全漏洞并进行实时有效的网络威胁态势评估，本文通过无监督生成网络模型

表 1 2 种无监督网络模型的优缺点分析

无监督网络模型	优点	缺点
VAE	与 AE 和 GAN 相比，VAE 在编码过程中通过对原始样本添加高斯噪声，使编码器能够更好地学习数据的先验分布，样本生成多样性更好	在度量生成样本与原始样本的相似度时，只能通过误差函数对数据元素间的相似度误差进行粗略计算，可能导致样本匹配的准确度降低
GAN	GAN 对生成样本与原始样本进行误差度量时，重构误差由参数化后的 GAN 的判别器决定，由于判别器为神经网络，其进行误差度量时是特征（语义）层面的，而不是简单的数据元素相似度误差计算，而 VAE 进行误差度量时，直接计算两者数据元素之间的 KL 散度和损失函数值，较简单，相比而言 GAN 的判别标准更高	生成器对真实样本分布的拟合不易收敛到一个较好的结果，且容易导致生成样本的多样性减少或者陷入模式坍塌（MC, model collapse）

对多源网络流量数据进行威胁分析。

本文提出一种网络威胁态势量化评估架构，如图 5 所示，主要包括评估数据集构建、数据预处理、多源数据特征选取、网络威胁测试和网络威胁态势评估 5 个部分。网络威胁量化评估步骤如下。

步骤 1　评估数据集构建。获取多源网络安全流量数据集作为评估数据源。

步骤 2　数据预处理。将原始数据进行数值化和特征规范处理，满足模型训练要求，提升数据利用率。

步骤 3　多源数据特征选取。针对多源网络安全流量数据进行特征选取，减小数据冗余度。

步骤 4　网络威胁测试。通过无监督威胁测试模型进行威胁测试，获取威胁发生概率。

步骤 5　网络威胁态势评估。根据步骤 4 计算得到的威胁发生概率，确定威胁严重度和威胁影响度，然后计算威胁态势值，对网络安全的整体态势进行分析评估。

3.2　评估数据集构建

3.2.1　评估数据集

为了保证 V-G 模型对多源网络流量数据进行特征解析时，所用数据具有高可利用性，本文选取不同领域的 4 个最具代表性的网络流量数据集作为评估数据集。4 个数据集的基本信息如表 2 所示。

表 2　　**流量数据集**

数据集	数据量/条	类别/种	流量数据来源
HTTP CSIC 2010	61 000	16	Web 应用程序
ADFA-LD	5 925	6	Linux 主机异常
ISOT	1 675 424	19	混合僵尸网络
UNSW-NB15	257 673	10	DDoS 匿名攻击

1) HTTP CSIC 2010 数据集是基于 Web 应用程序自动生成的正常和异常网络攻击流量数据集合，包含 36 000 条正常请求和超过 25 000 条异常请求，主要存在 3 种类型的异常请求，具体分为 16 个攻击类别。

① 静态攻击，是指尝试请求不允许的资源的攻击行为，这类资源包括过时文件、URL 重写中的会话 ID、默认文件、Web 程序配置文件等。

② 动态攻击，是指修改有效请求参数的攻击行为，这类攻击包括 SQL 注入、CRLF（carriage return/line feed）注入、参数篡改、信息收集、文件泄露、跨站点脚本、XSS（cross site scripting）漏洞、缓冲区溢出、恶意身份验证等。

③ 无意的非法请求，是指非恶意的不符合正常参数值结构的输入请求，例如，由字母组成的电话号码等。

2) ADFA-LD 数据集是基于 Linux 主机级入侵检测系统的网络流量数据集合，包含 5 925 条流量数据，主要分为以下 6 种攻击类别：Hydra-FTP、

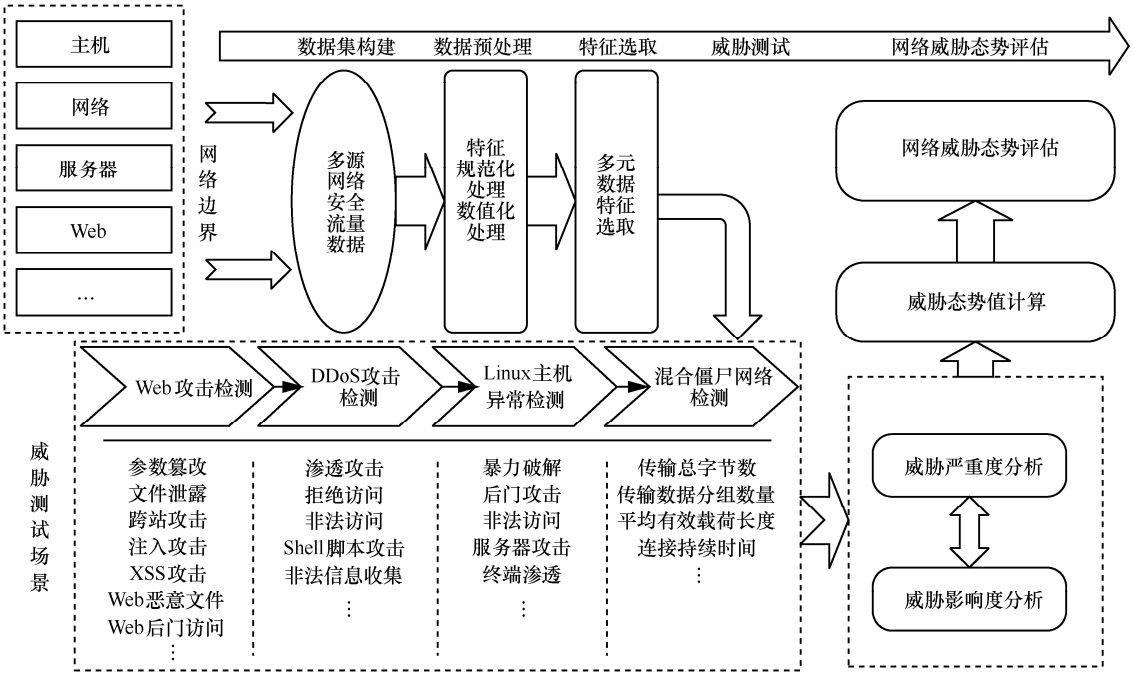


图 5　网络威胁态势评估架构

Hydra-SSH、Adduser、Java-Meterpreter、Meterpreter 和 Webshell。

3) 信息安全和对象技术（ISOT, information security and object technology）数据集由各种僵尸网络流量和正常网络数据流量组成，其包含 1 675 424 条流量数据，分为以下 19 种特征类别：BytesAB、BytesBA、Npackets、NpacketsAB、NpacketsBA、Duration、APL、DPS、Payload、PPS、BS、TBT、Flen、NNP、NSP、PSP、IPP、OPP 和 PV。

4) UNSW-NB15 数据集主要由 2007 年 DDoS（distributed denial of service）攻击中大约一小时的匿名流量跟踪数据组成，其包含 257 673 条流量数据，主要分为以下 9 种攻击类别：Fuzzers、Analysis、Backdoors、DoS、Exploits、Generic、Reconnaissance、Shellcode 和 Worms。

4 个数据集所包含的部分网络威胁态势指标如表 3 所示。

表 3 部分网络威胁态势指标

攻击类型	编号	态势指标
Web 攻击	1	文件泄露
	2	XSS 漏洞
	3	参数篡改
DDoS 匿名流量攻击	4	拒绝用户访问
	5	渗透攻击
	6	非法系统访问
主机异常检测流量	7	暴力破解
	8	后门攻击
	9	服务器攻击
混合僵尸网络流量	10	传输的数据分组数量
	11	平均有效载荷长度
	12	连接持续时间

表 3 列举了用于 V-G 模型威胁测试的部分威胁态势指标。针对本文研究中不存在的其他类型威胁指标，通过 V-G 模型同样能够进行有效测试，前提是获取包含这些攻击威胁的数据流量集合，因为威胁测试的前提是需要大量的网络流量数据作为基准数据进行模型训练。

3.2.2 数据预处理

数据预处理过程包括字符特征数值化和数值特征规范化两项操作。

1) 字符特征数值化

V-G 网络模型的训练需要数字特征向量作为输入项，针对多源网络安全评估数据集中的符号型数据需进行数值化处理，将所有符号型特征转换为一组有序数值特征。

具体地，通过独热编码（One-Hot）方式将 HTTP CSIC 2010 数据集中存在的 14 个 HTTP 请求特征类转化为数值向量，处理过程如下。

① 将 Protocal、userAgent、accept、accept-Encoding、pragma、cacheControl、acceptCharset 和 acceptLanguage 这 8 类特征数据转换为 0 和 1 之间的数值向量。

② 将 GET、POST 和 PUT 这 3 类 HTTP 请求数据分别转换为二进制特征向量（1,0,1）、（1,0,0）和（1,1,0）。

③将 JSP、GIF 和 PNG 等 Web 应用程序的 URL 扩展分别转换为二进制特征向量（1,1,1）、（0,1,1）和（0,1,0）。

同样地，UNSW-NB15 数据集经过数值化处理后，其每条原始数据的 42 维特征被转换为 196 维二进制数值向量。

2) 数值特征规范化

为消除量纲，抑制取值范围差异对网络训练的负面影响，同时提升训练效果，将经过数值化处理后的所有数值特征规范在[0,1]内，如式(1)所示。

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

其中，x^*为某类特征经过规范后的值，x 为初始特征值，$x_{\min}$为最小特征值，$x_{\max}$为最大特征值。

3.3 多源数据特征选取

为了降低数据集的冗余度，减小模型训练复杂度并提升模型的泛化能力，必须进行多源网络安全数据的特征选取。

由于 V-G 模型训练是一个多特征聚类过程，且所用训练数据集中数据具有多聚类结构特性，因此针对多源数据执行特征选取任务时，需满足以下 2 点要求：一是经过特征选取后的数据可以覆盖单一数据集中所有可能的聚类情况，二是能够最大程度地还原数据本身的聚类结构特性。

多簇特征选择（MCFS, multi-cluster feature selection）算法作为一种无监督特征选取算法，其进行特征选取任务时不需要依赖数据标签信息进行引导。为此，本文在多源数据特征选取时使用该算法。

特征选取过程包括以下步骤：构建 k–最近邻图、谱嵌入聚类分析、稀疏系数学习、计算 MCFS

分数和特征选择。特征选取过程如图 6 所示。

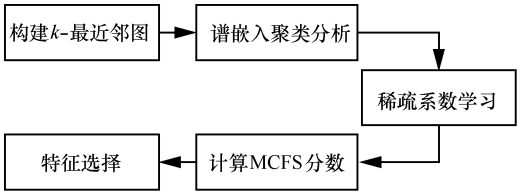


图 6　特征选取过程

1) 构建 k–最近邻图

对于具有 N 个顶点的图所对应的每个数据点 x_i，通过寻找 x_i 的 k 个最近邻点构建 k–最近邻图，以获取数据分布的局部几何结构特征和邻接权重矩阵 $\boldsymbol{W}$。本文使用热核权重法计算数据点间的邻接权重矩阵 $\boldsymbol{W}$，如式(2)所示。

$$W_{ij} = \mathrm{e}^{-\frac{|x_i - x_j|^2}{\sigma}} \tag{2}$$

其中，x_i 和 x_j 表示 k–最近邻图中的任意 2 个数据点，σ 为固定参数。

2) 谱嵌入聚类分析

定义一个对角矩阵 $\boldsymbol{D}$，$\boldsymbol{D}$ 的对角元素为 $D_{ii}=\sum_{i=1} w_{ij}$，通过计算拉普拉斯矩阵 $\boldsymbol{L}$ 的广义特征值可以获取数据流的平面嵌入结构，如式(3)所示。

$$\boldsymbol{L}\boldsymbol{h}_k = \lambda \boldsymbol{D}\boldsymbol{h}_k \tag{3}$$

其中，$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{W}$；$k$ 为数据的内在维度，其大小通常设为数据集的簇（聚类）的个数；λ 为拉普拉斯矩阵 $\boldsymbol{L}$ 的特征值。设 $H=\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_k\}$，它表示通过式(3)求得的最小广义特征值所对应的特征向量集合，H 的每一列表示任意一个数据点 x_i 的平面嵌入。

3) 稀疏系数学习

在获取数据点的平面嵌入 H 之后，为评估每个特征在其对应数据维度内（H 的每一列）的重要性和衡量每个特征对数据聚类的区分能力，给定 H 的列的任意一个嵌入 $\boldsymbol{h}_k$，MCFS 将其作为一个回归目标，目标函数如下。

$$\min_{\boldsymbol{a}_k} \left\| \boldsymbol{h}_k - \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{a}_k \right\|^2 + \kappa \left| \boldsymbol{a}_k \right| \tag{4}$$

其中，$|\boldsymbol{a}_k|$是一个 M 维向量，$\boldsymbol{Q}$ 是 $N\times M$ 维矩阵，κ是正则参数。为最小化该目标函数，定义 $\boldsymbol{a}_k$ 的 L1 范数[9]为

$$\left| \boldsymbol{a}_k \right| = \sum_{j=1}^{M} \left| a_{k,j} \right| \tag{5}$$

其中，$\boldsymbol{a}_k$ 包含用于近似计算 $\boldsymbol{h}_k$ 时不同特征的稀疏系数。根据 L1 范数的惩罚性质，当 κ 足够大时，$\boldsymbol{a}_k$ 的稀疏系数将逐渐缩小为零，此时可选择一个与 $\boldsymbol{h}_k$ 相关性最强的特征子集。

4) 计算 MCFS 分数

对于每个包含 k 个簇的数据集，使用步骤 3)的方法计算 k 个稀疏系数向量$\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_k\}\in \mathrm{R}^M$，其中每个非零元素 $\boldsymbol{a}_k$ 对应 d 个特征。为了能够从 k 稀疏系数向量中选取 d 个有效特征，对于每个特征 j，定义该特征的 MCFS 分数为

$$\mathrm{MCFS}(j) = \max_k \left| a_{k,j} \right| \tag{6}$$

其中，$a_{k,j}$ 表示向量 $\boldsymbol{a}_k$ 的第 j 个元素。

5) 特征选取

根据步骤 4)计算数据集中每一类特征的 MCFS 分数，将所有特征的 MCFS 分数降序排序，如式(7)所示。

$$S = \mathrm{sort}\left[\mathrm{MCFS}_1, \mathrm{MCFS}_2, \cdots, \mathrm{MCFS}_n\right] \tag{7}$$

通常 MCFS 分数越大表示特征越重要，因此根据排序结果，可从中选取前 d 个重要特征。

4　网络威胁态势量化评估

4.1　基于 V-G 模型的网络威胁测试方法

为对网络环境中可能出现的新的网络威胁进行实时检测，本文基于 V-G 网络进行网络威胁测试。基于 V-G 网络的威胁测试模型如图 7 所示。

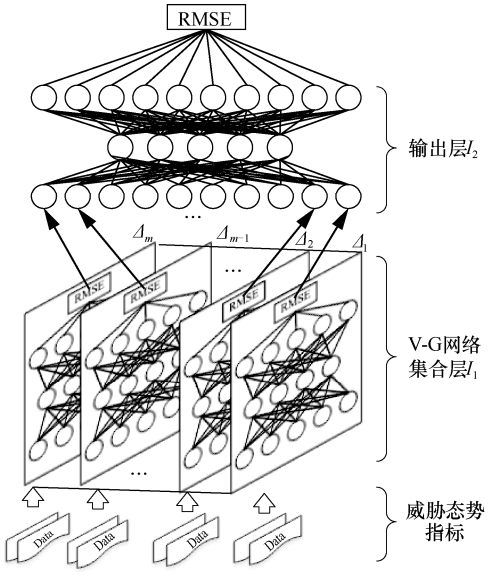


图 7　基于 V-G 网络的威胁测试模型

基于 V-G 网络的威胁测试过程主要包括 4 个阶段：网络集合层训练、网络参数优化、输出层重构误差训练和威胁测试。

为便于表述和分析，令 Δ 表示单个 V-G 网络，同时令 I_1 和 I_2 分别表示网络集合层和网络输出层，I_1 是由 m 个 Δ 组成的网络集合，I_2 是一个三层的变分自动编码器，具有 k 个输入和输出单元。

网络威胁测试过程如图 8 所示。无监督网络威胁测试过程详细步骤设计如下。

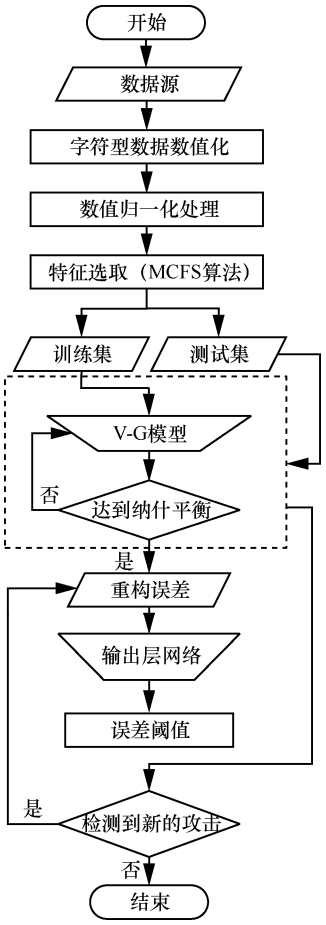


图 8 网络威胁测试过程

步骤 1 网络集合层的训练。将经过数据预处理和多源数据特征选取后的正常网络流量数据分批输入 I_1 进行训练，当训练达到纳什平衡状态时，结束训练。

步骤 2 网络参数优化。传统的梯度下降（GD, gradient descent）算法和高斯牛顿（GN, Gauss Newton）算法在参数调优过程中经常出现参数值陷入局部最优的情况，列文伯格（LM, Levenberg Marquardt）优化算法作为一种自适应优化算法，克服了上述 2 种算法的局限性，同时结合了 GD 和 GN 算法的优点，具备全局收敛能力。因此在网络集合训练过程中，为更有效地更新网络参数，本文使用 LM 优化算法代替 GD 和 GN 算法对 V-G 网络进行参数调优。LM 算法如算法 1 所示。

算法 1 LM 算法优化过程

输入 矢量函数 $\boldsymbol{f}$：$\mathrm{R}^m \to \mathrm{R}^n$（$n \geqslant m$），初始迭代向量 $\boldsymbol{x}_0 \in \mathrm{R}^m$，期望向量 $\boldsymbol{p} \in \mathrm{R}^n$

输出 使目标函数 $\|\boldsymbol{p}-\boldsymbol{f}(\boldsymbol{x})\|^2$ 最小化的向量 $\boldsymbol{x}^*$

begin

$l=0$; $v=2$; $\boldsymbol{x}=\boldsymbol{x}_0$; //$l$ 为初始迭代次数，v 为向量维数

$\boldsymbol{D}=\boldsymbol{J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{J}(\boldsymbol{x})$; //$\boldsymbol{D}$ 表示使用 Jacobian 矩阵近似计算 Hessian 矩阵

$\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{p}-\boldsymbol{f}(\boldsymbol{x})$; $\boldsymbol{g}=\boldsymbol{J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{x})$; //$\boldsymbol{F}(\boldsymbol{x})$为误差向量，$\boldsymbol{g}$ 为近似计算 Hessian 矩阵的更新

$\mu=\tau * \max\limits_{i=1,\cdots,m}(D_{ii})$; //$\mu$ 为阻尼因子

stop = ($\|\boldsymbol{g}\|_\infty \leqslant \eta_1$); //$\eta_1$ 为迭代终止临界值

while（not stop）and（$l \leqslant l_{\max}$） //$l_{\max}$ 表示最大迭代次数

$l=l+1$;

repeat

solve $(\boldsymbol{D}+\mu\boldsymbol{I})\,\boldsymbol{h}_x=\boldsymbol{g}$; //通过方程求解 $\boldsymbol{h}_x$

if（$\|\boldsymbol{h}_x\| \leqslant \eta_2\|\boldsymbol{x}\|$）

stop = true;

else

$\boldsymbol{x}_{\mathrm{new}}=\boldsymbol{x}+\boldsymbol{h}_x$;

$$\delta=\frac{\left\|F(x)\right\|^2-\left\|p-f(x_{\mathrm{new}})\right\|^2}{h_x^{\mathrm{T}}(\boldsymbol{h}_x+g)};$$

//δ 为一个系数，控制 μ 的更新

if $\delta>0$

$\boldsymbol{x}=\boldsymbol{x}_{\mathrm{new}}$;

$\boldsymbol{D}=\boldsymbol{J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{J}(\boldsymbol{x})$; $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{p}-\boldsymbol{f}(\boldsymbol{x})$; $\boldsymbol{g}=\boldsymbol{J}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{x})$;

stop = ($\|\boldsymbol{g}\|_\infty \leqslant \eta_1$ or ($\|\boldsymbol{F}(\boldsymbol{x})\|^2 \leqslant \eta_3$));

$$\mu=\mu*\max\left(\frac{1}{3},\,1-(2\delta-1)^3\right);$$

$v=2$;

else

$\mu=\mu*v$; $v=2*v$;

end if

end if

until ($\delta > 0$) or (stop)

end while

$\boldsymbol{x}^* = \boldsymbol{x}$;

end

步骤 3　输出层重构误差训练。输出层网络 I_2 的输入项来自 I_1 中每个相应子网络训练输出的 0-1 归一化重构误差值。由 I_1 和 I_2 输出的重构误差值均由均方根误差（RMSE, root mean square error）函数计算得到，如式(8)所示。

$$\mathrm{RMSE}(\boldsymbol{x},\boldsymbol{y}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{8}$$

其中，$\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别表示输入样本向量和生成样本向量，n 为输入向量的维数。

由 I_1 输出的训练误差集合 ξ^*可表示为

$$\xi^* = \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_m\} \tag{9}$$

将 ξ^*作为 I_2 的输入项进行重构误差训练，最终通过 RMSE 函数统计训练阈值 η。

步骤 4　威胁测试。随机从包含异常流量数据的测试数据集 $X^{(2)}$中选取 m 组相同数量的测试样本集合 v= {$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$}，将其作为测试数据集进行威胁测试。

每次测试由 I_1 输出的测试误差 λ 为

$$\lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_m\} \tag{10}$$

4.2　网络威胁态势量化评估

在本文研究中，网络威胁态势量化评估结果由影响网络安全的 2 个关键因素——威胁严重度和威胁影响度确定。

1) 威胁严重度

本文使用无监督网络模型对多源网络流量数据进行特征解析，在执行威胁测试任务后，根据威胁测试结果，将每次测试时获取的归一化测试误差值 λ 作为威胁发生概率（TP, threat probability），即

$$\mathrm{TP}_i = \lambda_i \tag{11}$$

为通过威胁发生概率确定威胁严重度等级，本文参考《国家突发公共事件总体应急预案》并结合 Snort 手册的攻击分类，对网络威胁态势严重度进行等级分类，将威胁严重度划分为安全、低危、中危、高危和超危 5 个等级，分别对应 5 个威胁发生概率区间：0.00~0.20 、 0.21~0.40 、 0.41~0.60 、 0.61~0.80 和 0.81~1.00，如表 4 所示。

表 4　威胁严重度等级划分

威胁严重度	概率区间	说明
安全	0.00~0.20	网络运行稳定正常，没有超出正常认知的恶意行为或具有严重威胁的安全漏洞被发现
低危	0.21~0.40	网络运行受到轻微影响，有少量威胁的安全漏洞被发现
中危	0.41~0.60	网络运行受到一定影响，有较高威胁级别的安全漏洞被发现
高危	0.61~0.80	网络运行受到较大影响，Web 攻击、非法访问等活动增加，多种威胁类型被发现
超危	0.81~1.00	网络运行受到严重影响，有大量异常攻击行为和威胁级别很高的安全漏洞被发现

2) 威胁影响度

为对网络威胁发生的影响程度进行划分，参照通用漏洞评分系统（CVSS, common vulnerability scoring system）[10]制定威胁影响度等级分类，如表 5 所示。

表 5　威胁影响度等级划分

影响度	概率区间	影响指标		
		机密性（C）	完整性（I）	可用性（A）
无影响	0.00~0.40	0	0	0
低影响	0.41~0.80	0.22	0.22	0.22
高影响	0.80~1.00	0.56	0.56	0.56

威胁影响度（TI, threat impact）定义为

$$\mathrm{TI} = \mathrm{lb}\frac{\chi_1 2^C + \chi_2 2^I + \chi_3 2^A}{3} \tag{12}$$

其中，C、I、A 分别表示机密性、完整性和可用性 3 个威胁影响度指标，χ_1、χ_2 和 χ_3 分别对应 3 个威胁影响度指标的权重。

威胁态势值（TSV, threat situation value）表示由威胁发生概率和威胁影响度 2 个威胁影响因素决定，定义为

$$\mathrm{TSV} = \frac{1}{n}\sum_{i=1}^{n}[\mathrm{TP}_i][\mathrm{TI}_i] \tag{13}$$

5　实验与结果

5.1　实验环境

基于 V-G 网络的训练和测试过程均在 Ubuntu 系统上进行，使用 Python 语言编程实现算法。实验的硬件环境为 Intel Core i7-7700 HQ 处理器，8 GB

RAM，显卡为 GTX1050，内存为 16 GB。

5.2 网络威胁测试结果分析

5.2.1 网络训练

在网络训练阶段，使用 AE、VAE、GAN 和 V-G 这 4 个网络分别组成无监督威胁测试模型，在不同的网络集合层数下分别进行模型训练。4 个模型均使用相同的网络训练参数，所用训练数据为同一个正常网络流量数据集。

在同一训练集和不同的网络集合层数下，4 个威胁测试模型在模型训练阶段输出的训练异常阈值 η 如图 9 所示。

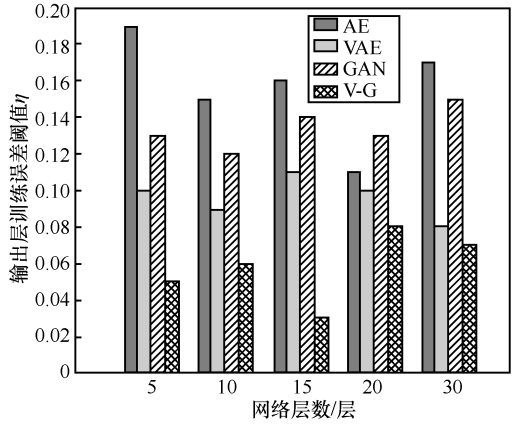


图 9 不同网络模型的训练阈值 η

由图 9 可见，与其他 3 类模型相比，基于 V-G 的威胁测试模型在训练阶段获取的异常阈值 η 最小，表明与其他 3 种模型相比，V-G 对原始数据的重构能力更好，更有利于提升威胁测试准确度。当网络集合层数达到 15 层时，基于 V-G 的威胁测试模型的 η 达到最小值。

当网络集合层数为 15 层时，V-G 网络在训练和测试过程中的迭代次数与误差的变化趋势如图 10 所示。

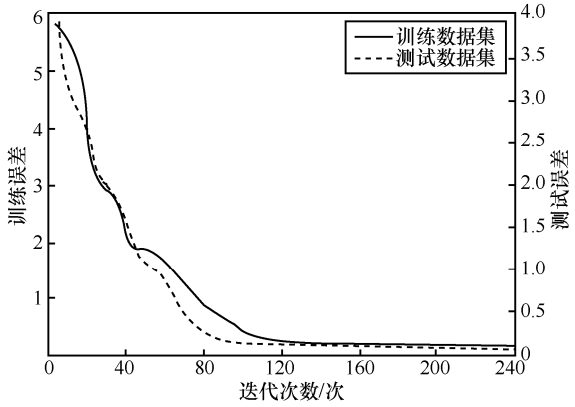


图 10 迭代次数与误差的关系

由图 10 可见，V-G 网络的训练误差曲线和测试误差曲线拟合度较好，说明模型达到了较好的收敛效果，能够更好地保证量化评估准确度。

在网络训练过程中，使用 GD、NM、GN 和 LM 这 4 种优化算法分别对 V-G 网络的参数进行优化，4 种算法优化过程的收敛情况如表 6 所示。

表 6 不同优化算法的收敛情况

优化算法	迭代次数/次	时间/s	RMSE
GD	220	348	0.35
NM	210	366	0.37
GN	200	323	0.32
LM	240	341	0.08

由表 6 可见，当模型达到收敛状态时，LM 算法相较于其他 3 种算法虽然迭代次数较多，耗时较长，但所得均方根误差值最小，因此模型收敛效果最好，能够有效提高威胁测试精确度。

5.2.2 网络威胁测试

在威胁测试阶段，当网络集合层数为 15 层时，使用 AE、VAE、GAN 和 V-G 这 4 个无监督威胁测试模型分别进行威胁测试实验，为保证测试实验结果具有可比性，使用同一测试数据集中进行实验，其中 10 组实验的归一化测试误差值 β 如图 11 所示。

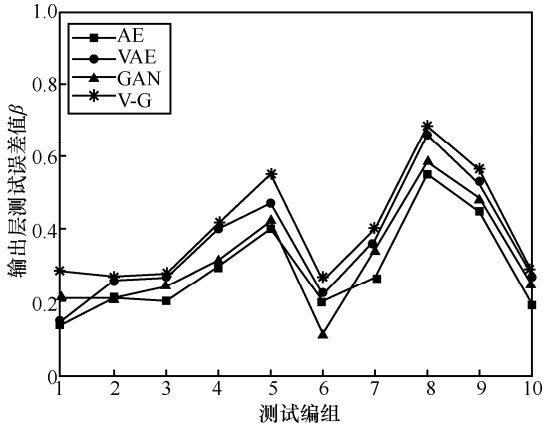


图 11 4 个模型的威胁测试结果

由图 11 可见，在同一样本测试集下，当网络集合层数达到 15 层时，与其他 3 种模型相比，基于 V-G 网络的测试模型的测试误差值 β 最大，这说明其对网络威胁的检测能力更突出。

5.3 网络威胁态势量化评估结果分析

在得到每组测试的输出层归一化误差值 β 后，

对照表 4 和表 5 分别确定威胁严重度和威胁影响度。

当 V-G 的网络集合层数为 15 层时，选取其中 10 组威胁测试实验的威胁发生概率，得到的威胁态势严重度和威胁影响度评估结果如表 7 所示。

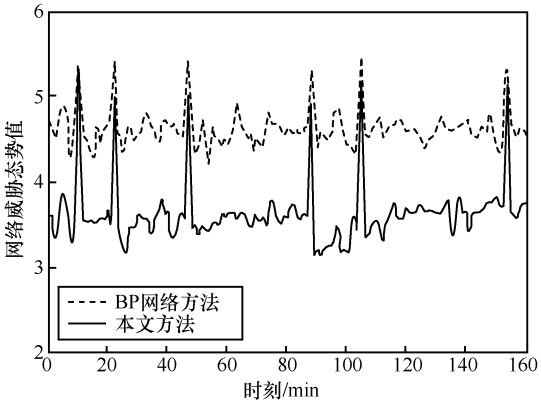
表 7　威胁严重度和威胁影响度评估结果

编号	威胁概率	威胁严重度	威胁影响度
1	0.187	安全	无影响
2	0.275	低危	无影响
3	0.238	低危	无影响
4	0.426	中危	低影响
5	0.557	中危	低影响
6	0.262	低危	无影响
7	0.358	低危	无影响
8	0.685	高危	高影响
9	0.504	中危	低影响
10	0.281	低危	无影响

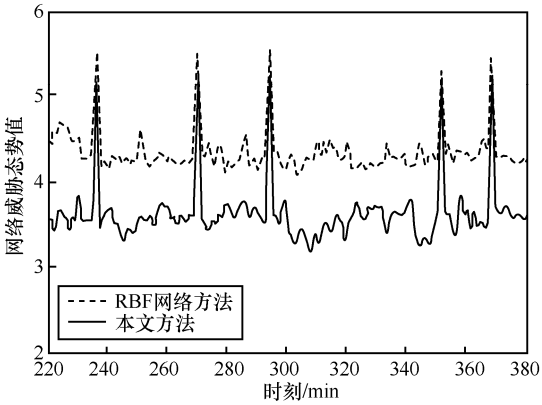
为了保证评估结果的客观性和真实性，首先根据威胁测试过程中获取的威胁发生概率和威胁影响度评估结果计算威胁态势值，然后使用 BP（back propagation）网络方法[11]和 RBF（radial basis function）网络方法[12]分别计算威胁态势值并与本文方法的威胁态势值计算结果进行比较，在某 2 个时间段内，由 3 种方法得到的威胁态势值如图 12 所示。

由图 12(a)可见，在时刻分别为 9 min、22 min、47 min、89 min、108 min 和 153 min 时，威胁态势值变化幅度较大，表明在这几个时刻网络遭受威胁的严重程度较高，且网络正遭受多种类型攻击。通过数据分析发现，在网络遭受严重威胁的 6 个时刻，与 BP 网络方法相比，本文方法对网络威胁的表征能力更强。例如，在第 9 min 时，基于 BP 网络方法的威胁态势值由 4.72 变为 5.39，而本文方法得到的威胁态势值则由 3.68 变为 5.40。相比之下，本文方法得到的威胁态势值的变化幅度更大。同样地，由图 12(b)可见，在网络遭受攻击的 5 个时刻，本文方法相较于 RBF 网络方法具有更直观的威胁表征效果。

某一评估阶段，3 种方法完成量化评估任务所耗费时间的对比情况如图 13 所示。



(a) BP网络方法与本文方法对比



(b) RBF网络方法与本文方法对比

图 12　威胁态势值对比

由图 13 可知，在相同数据量下，与 BP 网络方法和 RBF 网络方法相比，本文方法耗时最少，这表明本文方法的评估效率更高。

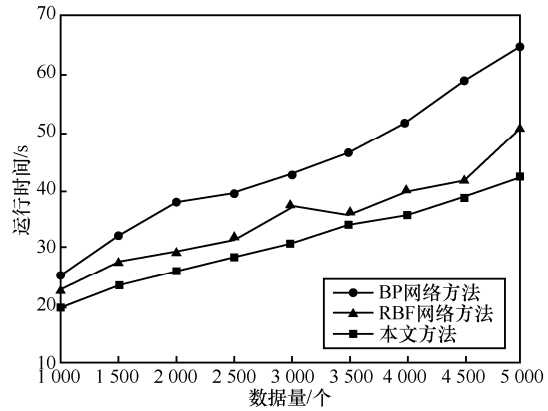


图 13　3 种方法的评估耗时对比

6　结束语

为克服传统的基于监督式特征学习的网络威胁态势量化评估方法需依靠数据标签进行建模的局限性，本文提出一种基于无监督多源数据特征解析的网络威胁态势评估方法，采用无监督学习方法

对多源网络流量数据进行特征解析，根据量化后的网络威胁态势影响因子值计算威胁态势值，完成对网络威胁态势的量化评估任务。实验结果表明，本文方法对网络威胁的表征能力较强，同时拥有较出色的网络威胁态势评估效果。

参考文献：

[1] YANG M, JIANG R, GAO T L, et al. Research on cloud computing security risk assessment based on information entropy and Markov chain[J]. International Journal of Network Security, 2018, 20(4): 664-673.

[2] WANG H, CHEN Z, FENG X, et al. Research on network security situation assessment and quantification method based on analytic hierarchy process[J]. Wireless Personal Communications, 2018, 102(2): 1401-1420.

[3] SALLAM H. Cyber security risk assessment using multi fuzzy inference system[J]. International Journal of Engineering and Innovative Technology, 2015, 4(8): 13-19.

[4] 文志诚, 陈志刚, 唐军. 基于信息融合的网络安全态势量化评估方法[J]. 北京航空航天大学学报, 2016, 42(8): 1593-1602.
WEN Z C, CHEN Z G, TANG J. Network security situation quantitative evaluation method based on information fusion[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(8): 1593-1602.

[5] FENG W, WU Y, FAN Y. A new method for the prediction of network security situations based on recurrent neural network with gated recurrent unit[J]. International Journal of Intelligent Computing and Cybernetics, 2018, 11(4): 511-525.

[6] HE F, ZHANG Y, LIU D, et al. Mixed wavelet-based neural network model for cyber security situation prediction using modwt and hurst exponent analysis[C]// International Conference on Network and System Security. Springer-Verlag, 2017: 99-111.

[7] DOERSCH C. Tutorial on variational autoencoders[J]. arXiv Preprint arXiv: 1606.05908, 2016.

[8] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// The 27th International Conference on Neural Information Processing Systems. MIT Press, 2014: 1-9.

[9] LAENG E, MORPURGO C. An uncertainty inequality involving L1 norms[J]. Proceedings of the American Mathematical Society, 1999, 127(12): 3565-3572.

[10] MELL P, SCARFONE K, ROMANOSKY S. Common vulnerability scoring system[J]. IEEE Security and Privacy Magazine, 2012, 4(6): 85-89.

[11] 唐成华, 余顺争. 一种基于似然 BP 的网络安全态势预测方法[J]. 计算机科学, 2009, 36(11): 97-100.
TANG C H, YU S Z. A network security situation prediction method based on likelihood BP[J]. Computer Science, 2009, 36(11): 97-100.

[12] 赖智全. 基于混合优化 RBF 神经网络的网络安全态势预测模型[D]. 兰州: 兰州大学, 2017.
LAI Z Q. Network security situation prediction model based on hybrid optimization RBF neural network[D]. Lanzhou: Lanzhou University, 2017.

[作者简介]

杨宏宇（1969− ），男，吉林长春人，博士，中国民航大学教授，主要研究方向为网络信息安全。

王峰岩（1993− ），男，河南南阳人，中国民航大学硕士生，主要研究方向为网络信息安全。