文章编号:1001-9081(2016)S2-0049-05

基于信息增益特征选择的网络异常检测模型

刘汝隽[1*],贾 斌[2],辛 阳[1]
(1.北京邮电大学 计算机学院,北京 100876; 2.北京邮电大学 网络技术研究院,北京 100876)
(* 通信作者电子邮箱 liurujun1211@126.com)

摘 要:针对网络异常检测中数据的高维特征会影响检测率和实时检测效率等问题,提出了一种基于信息增益特征选择的网络异常检测模型。首先,预处理器将网络流量数据规范化;其次,基于信息增益降维方法的特征选择器选取重要特征,降低数据集的维度;最后,随机森林分类器经过训练和预测得到检测结果。实验中,该模型能够将随机森林分类器的检测率提高0.2%,将检测时间平均缩短19%;在检测率上优于K近邻算法,在误报率、阳性似然比和约登指数方面均优于K近邻和AdaBoost算法。实验结果表明,所提模型能够有效提高检测率,缩短检测时间。

关键词:网络异常检测;信息增益;特征选择;分类;随机森林
中图分类号:TP309 **文献标志码**:A

Network anomaly detection model based on information gain feature selection

LIU Rujun[1*], JIA Bin[2], XIN Yang[1]
(1. *School of Computer Science and Technology, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
2. *Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*)

Abstract: In order to deal with the issue that high dimensional feature data in network anomaly detection will reduce detection rate and real-time detection efficiency, a network anomaly detection model based on information gain feature selection was proposed. Firstly, network flow data was standardized in preprocessor; secondly, important features were selected by feature selection container based on iinformation gain dimension-reducing method to reduce the dimension of the dataset; finally, detection results were predicted by Random Forest (RF) classifier. In the experiments, the detection rate of Random Forest classifier was increased by 0.2% and the detection time was shortened by 19% in average; the model performed better than KNN (K-Nearest Neighbor) in detection rate and better than KNN and AdaBoost algorithms in false positive rate, positive likelihood ratio and Youden index. The experimental results show that the proposed model can improve detection rate, and reduce detection time.

Key words: network anomaly detection; Information Gain (IG); feature selection; classification; Random Forest (RF)

0 引言

互联网的迅速发展在推进网络技术进步的同时也导致网络安全事件频繁发生,传统的计算机安全理论已不能适应动态变化的、多维互联的网络环境[1]。随着云时代的来临和大数据技术的发展,网络异常事件对网络设备造成的影响更大、范围更广。作为网络异常行为中最突出的问题,网络攻击主要通过侵入主机来达到恶意破坏的目的。因此,通过分析网络节点的流量数据来分析是否存在入侵现象,成为识别网络安全事件的关键步骤[2]。

对此,研究者们对网络异常检测技术进行了大量的研究。文献[3]提出一种基于活跃熵的检测方法,无需花费大量时间对数据特征进行训练,大大提高检测效率。文献[4]提出一种基于改进极限学习机器的检测算法,检测率达到98%,误报率仅为1.74%。文献[5]提出一种基于改进支持向量机的检测算法,缩短数据的训练时间。作为数据挖掘经典算法,随机森林、K近邻和AdaBoost算法被广泛应用于网络异常检测。文献[6]使用随机森林算法对网络流量进行分类,并达到较高准确率。文献[7]使用主成分分析算法进行特征选择,采用随机森林算法进行异常检测,结果表明随机森林算法在检测率、误报率和时间复杂度上都有良好表现。文献[8]使用K近邻作为数据训练算法,达到提高准确率的目的。文献[9]将通过归一化处理后的数据使用K近邻算法进行分类,结果表明误报率只有2%。文献[10]采用AdaBoost算法对网络异常流量进行分类检测,其误报率和漏报率分别为0.3%和0.4%。文献[11]对AdaBoost算法进行改进,将检测率提高4%,误报率和漏报率分别降低4.6%和0.7%。

尽管现有的研究成果能够满足网络异常检测的基本需要,但高维数据对检测效果的影响方面的研究还不够充分,对于检测算法时间效率的定量分析还有待加强。

针对网络异常检测中数据的高维特征会降低分类检测的检测率和实时检测效率的问题,本文提出一种基于信息增益特征选择方法的异常检测分类模型,通过特征选择对随机森林分类算法进行改进,提高其检测率,缩短检测时间。

1 信息增益特征选择

1.1 相关背景

信息增益又名互信息,是一种基于信息论的特征选择方

收稿日期:2016-05-09;修回日期:2016-05-31。 基金项目:国家863计划项目(2015AA017201)。

作者简介:刘汝隽(1991—),女,河北承德人,硕士研究生,主要研究方向:信息安全、计算机网络安全; 贾斌(1982—),男,山东青岛人,博士研究生,主要研究方向:计算机网络安全、数据挖掘; 辛阳(1977—),男,山东海阳人,副教授,博士,主要研究方向:移动通信网络安全、计算机网络安全。

法。信息论由香农提出，它采用划分数据集的方式将杂乱无章的数据规则化，信息增益是规则化前后数据集发生的变化[12]。

数据集样本中属性的信息增益越大，其包含的信息量也越大。也就是说，在特征选择时应计算各个属性的信息增益，具有最高信息增益值的属性是给定集合中具有最高区分度的属性[13]。

1.2 方法描述

信息增益特征选择方法的定义如下：

定义1 数据集$\{x_1, x_2, \cdots, x_M\}$[14]，$X$是其中一个随机数，$p_i := P(x_i)$是选择某种分类方式的概率。$x_i$的信息期望值为：

$$L(x_i) = -\text{lb}\ p(x_i) \tag{1}$$

定义2 熵是信息论中广泛使用的度量标准，它是所有可能值的信息期望值：

$$Entropy(S) = -\sum_{i=1}^{n} p(x_i)\ \text{lb}\ p(x_i) \tag{2}$$

定义3 信息增益是指某属性分割样例导致的期望熵降低，信息增益由熵计算得来：

$$Gain(S, A) = Entropy(S) - \sum_{v \in V(A)} \frac{|S_v|}{S} Entropy(S_v) \tag{3}$$

其中：A是属性，$V(A)$是属性A的值域，S是样本的集合，S_v是S中在属性A上值等于v的样本集合。

2 检测模型

本文提出一种基于信息增益特征选择的异常检测分类模型，如图1所示。该模型在一定程度上能够提高检测率、缩短检测时间。

模型主要包括预处理器、特征选择器和分类器三部分，其中特征选择器和分类器是整个模型的核心模块。模型的工作流程如下：

1) 预处理器对数据集进行规范化，提取重要的异常类型，并将离散型数据转化为二进制数值。

2) 特征选择器将经过预处理的数据使用信息增益算法进行降维处理，去掉冗余数据，形成新的数据集。

3) 经过特征选择的数据进入分类器，经过训练和预测，得到分类结果。

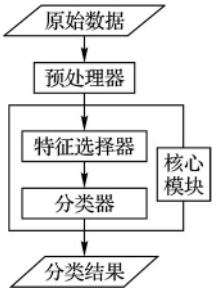

图1 模型流程

2.1 预处理器

预处理器的主要功能为数据规范化，包括以下几点内容：

1) 由于子类异常类型较多，将子类异常类型转化为父类异常类型。

2) 将离散型字符数据转换为二进制数，使数据结构统一化，便于计算。

3) 将类型较多、所占比重较低的异常类型合并为一类，简化数据。

2.2 核心模块

本文模型的核心部分主要包括特征选择器和分类器，数据经过预处理后，首先进入特征选择器，降维形成新的数据集，再进入分类器进行训练和分类。

2.2.1 特征选择器

特征选择器主要进行两项工作：特征子集产生和特征子集评估。特征子集产生是通过特征选择方法，选出特征子集；特征子集评估是特征子集的优化过程，使用评估函数得到最优子集。

2.2.2 分类器

本文模型采用随机森林[15]分类器进行分类和预测，随机森林算法主要为决策树而设计，该算法结合多棵树的预测结果，每棵树的产生都基于随机向量。随机森林算法在处理多特征的海量数据时效率较高[16]。图2是随机森林算法的工作流程，其中：D为原始数据；$D_1, D_2, \cdots, D_t$为随机向量；$T_1, T_2, \cdots, T_t$为多棵决策树；T^*为组合决策树。

随机森林预测算法 Classify($node$, V) 伪码描述如下：

输入 决策树节点$node$、测试数据V。

输出 预测标签$label$。

```
1)  if (node = leaf)
2)      return label
3)  else
4)      j = node. attribute
5)      if j is categorical
6)          v = V_j
7)          childv = node of v
8)          return Classify(childv, V)
9)      else if (j is real-valued)
10)         t = node. threshold
11)         if (V_j < t)
12)             childLO = leftNode
13)             return Classify(childLO, V)
14)         else
15)             childHI = rightNode
16)             return Classify (childHI, V)
17)         end
18)     end
19) end
```

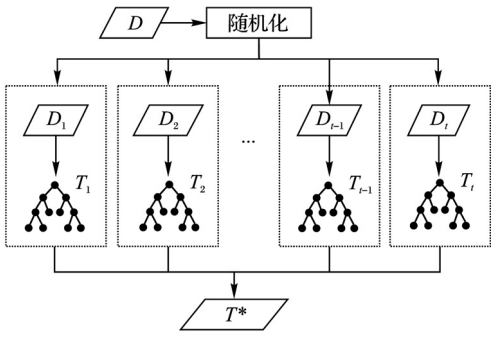

图2 随机森林算法流程

2.2.3 核心模块工作流程

特征选择器和分类器是本文模型的核心，将经过预处理的数据依次经过特征选择器和分类器即可得到分类结果。核心部分工作流程如下：

1) 经过预处理的数据进入特征选择器，特征选择器为每一个特征计算信息熵。

2) 根据每个特征的信息熵，得到每一种信息增益的划分结果。

3) 将划分结果排序之后作降维处理,得到有效数据。

4) 有效数据进入分类器。如果该数据是训练数据,就进入分类器的训练模块; 如果数据为测试数据,就进入分类器的测试模块。

5) 训练数据经过训练之后得到有效结果,测试数据根据该结果进行预测。

6) 得到预测结果,即完成一次分类过程。

3　实验与结果分析

实验部分主要包括实验环境与数据、实验评估指标和实验过程与结果分析三部分内容。

3.1　实验环境与数据

本文以 Windows 服务器、Matlab 2015b 作为实验平台,使用 1999 年数据挖掘与知识发现(Data Mining and Knowledge Discovery,KDD) 竞赛数据作为实验数据集。该数据集由林肯实验室采集得到,包括训练数据和测试数据,被广泛应用于网络异常检测。该数据集共有 41 种特征,实验中使用该数据集的一个子集。为保证实验结果的准确性和数据抽样的随机性,实验使用随机函数在一种正常类型和四种异常类型中分别随机获取与原数据集同样比例的数据,子集包括 4 940 条训练数据和 3 110 条测试数据。数据集各类型比例如表 1。

表 1　数据集类型

类别	类别描述	具体类型	训练数据占总数据比例/%	测试数据占总数据比例/%
Normal	正常流量	无	19.70	19.00
DoS	拒绝服务攻击	ping-of-death 等	79.00	74.00
R2L	来自远程主机的未授权访问	guessing password	0.22	5.20
U2R	未授权的本地超级用户访问	buffer overflow attacks	0.01	0.07
PROBING	端口监视或扫描	port-scan, ping-sweep	0.83	1.30

3.2　评估指标

本文采用异常检测常用的四种指标来评估实验,四类指标分别为检测率、误报率、阳性似然比和约登指数; 同时使用时间效率作为检测实验效果的指标之一。检测指标描述如下: I 和 $\neg I$ 分别代表正常行为和异常行为; A 和 $\neg A$ 分别代表被发现的异常行为和未被发现的异常行为。检测率 R_d(Detection Rate) 是被发现的异常行为与正常行为的比率; 误报率 R_{f+}(False Positive Rate) 为未被发现的异常行为与异常行为的比率; 阳性似然比 P 是检测率与误报率的比率; 约登指数[17] Q 是检测率与误报率之差。阳性似然比的取值范围为 0 到正无穷,数值 1 代表预测结果约等于随机值,数值越接近正无穷代表预测结果越好。约登指数取值范围为 $-1 \sim 1$,数值越高代表预测结果越好,数值 0 表示预测值约等于随机值; 相反地,数值 -1 代表预测值比随机值差。四项评估指标的计算公式如下:

$$R_d = A/I \tag{4}$$

$$R_{f+} = \neg A/I \tag{5}$$

$$P = R_d/R_{f+} \tag{6}$$

$$Q = R_d - R_{f+} \tag{7}$$

3.3　实验过程与结果分析

3.3.1　数据预处理

数据集由一种正常流量和四种异常流量组成,四种异常流量类型分别为 DoS、R2L、U2R 和 PROBING。每一种异常类型包括多个子类型,原数据集中每条数据均被标记为子类型。因此,首先将子类型转化为父类型。其次,由于特征 2 和 3 是离散型数据,因此需要将其进行规范化。特征 2 包括 TCP、UDP 和 ICMP 三种类型,使用二进制数 001,010 和 100 代替。特征 3 类型较多,超过 60 种,因此先对该特征的类型进行统计和排序。前三种类型的数据所占比率和超过 90%,其他类型所占比例较低,因此可将其他类型归为一类。特征 3 的规范化结果为: ecr_i、private、http 和其他类型,分别用 0001、0010、0100 和 1000 表示,所占比例分别为 56.97%、22.45%、13.01% 和 7.58%。

3.3.2　特征选择

使用信息增益算法对预处理后的数据集进行特征选择,结果表明: 前 16 种特征的重要性较高,其余特征的重要性均低于 0.1,因此本文选取前 16 种特征作为新的数据集。特征名称和重要性排名如图 3,选取的前 16 种特征的特征序号和名称如表 2。

3.3.3　实验结果分析

实验从两个角度对本文模型的检测效果进行分析和比较。首先将本文模型和随机森林分类器进行对比,其次将本文模型和 K 近邻、AdaBoost 分类器进行对比。

1) 从检测率和时间效率两个方面将本文模型与随机森林分类器的分类效果进行对比。实验结果如图 4 和表 3 ~ 4。

实验表明,本文方法比随机森林分类器的分类效果更好,如图 4,在选取的 8 个阈值中,本文方法的检测率均高于随机森林分类器,并且在选取阈值为 50 时,检测率最高能够提高 0.25%。说明经过信息增益特征选择降维后,能够对随机森林分类器的决策树达到剪枝效果,提高其分类效率。

表 2　特征重要性排名和名称

特征重要性	特征名称	特征重要性	特征名称
1	src_bytes	9	protocol_type
2	count	10	dst_host_srv_diff_host_rate
3	service	11	dst_host_srv_count
4	dst_bytes	12	dst_host_diff_srv_rate
5	srv_count	13	dst_host_same_srv_rate
6	logged_in	14	srv_diff_host_rate
7	dst_host_same_src_port_rate	15	diff_srv_rate
8	dst_host_count	16	same_srv_rate

表 3 和 4 分别计算了使用本文方法和使用随机森林分类器在训练过程和预测过程的使用时间和时间缩短比率。

表 3　训练过程时间效率

阈值	训练时间/s		时间缩短比率/%
	本文方法	随机森林分类器	
25	0.54	0.77	30.24
50	1.07	1.35	20.70
75	1.49	1.70	12.31
100	1.84	2.26	18.57
125	2.30	2.78	17.06
150	2.73	3.31	17.56
175	3.17	3.63	12.66
200	3.45	5.19	33.52

结果表明,本文方法在训练过程和预测结果过程的时间

均少于随机森林分类器，其平均时间缩短比率分别为 20% 和 17%。因此，本文方法在检测时间上较随机森林方法有很大改进。

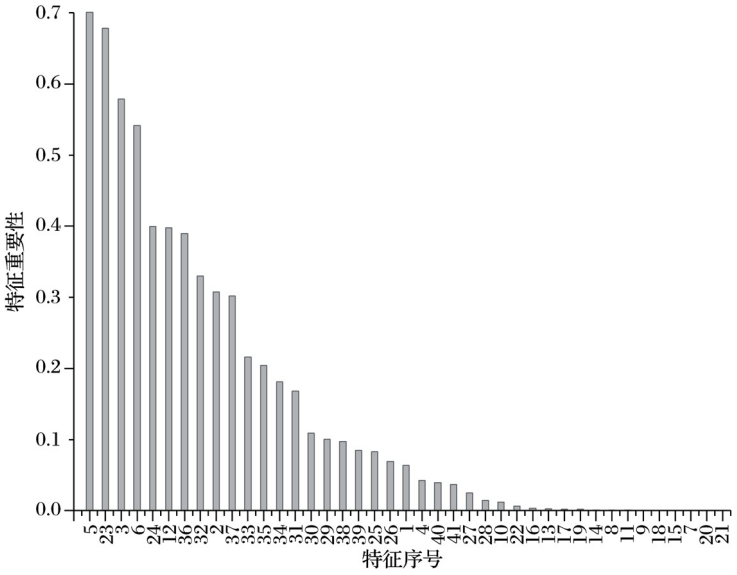

图 3 特征重要性

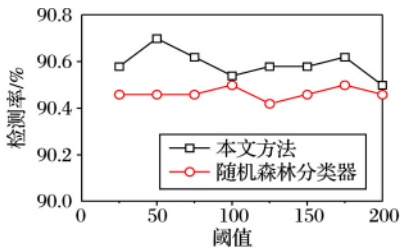

图 4 检测率对比

表 4 预测过程时间效率

阈值	预测时间/s		时间缩短比率/%
	本文方法	随机森林分类器	
25	0.17	0.21	18.03
50	0.23	0.40	41.86
75	0.43	0.54	19.72
100	0.57	0.65	13.45
125	0.72	0.81	11.01
150	0.84	0.97	13.83
175	1.13	1.25	10.06
200	1.27	1.39	8.61

2) 将本文模型与 K 近邻和 AdaBoost 分类器进行实验和比较，比较结果如图 5 ~ 8。

图 5 为不同阈值下三种算法的检测率比较，三种分类算法均能达到超过 90% 的分类结果。本文方法和 AdaBoost 算法受阈值影响较小，无论选择何种阈值均能达到较好的准确率；K 近邻算法准确率取决于阈值大小，但在选择适当的阈值时也能有良好表现。

图 6 为三种算法的误报率比较，在该指标中，只有本文方法达到较好结果，其误报率只有 0.2%，K 近邻算法误报率达到 2%，AdaBoost 算法误报率超过 10%。

图 7 为三种算法的阳性似然比指标比较，结果表明，本文方法在该指标的结果远远优于其他两种算法；三种算法的约登指数比较如图 8，本文方法的结果依然优于其他两类算法，并且几乎不受阈值影响，较为稳定。

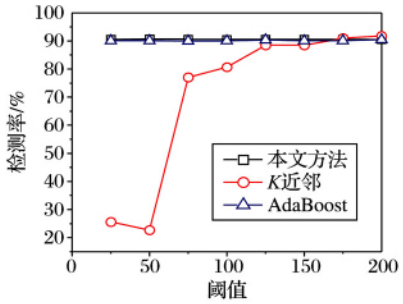

图 5 三种算法检测率比较

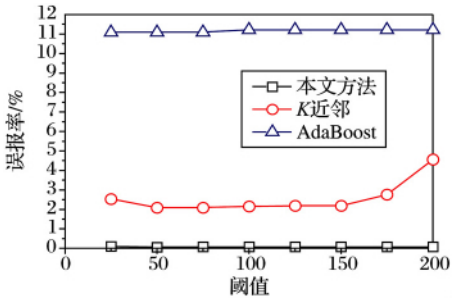

图 6 三种算法误报率比较

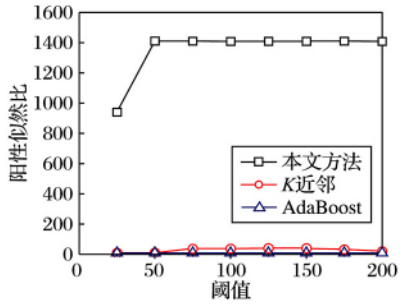

图 7 三种算法阳性似然比的比较

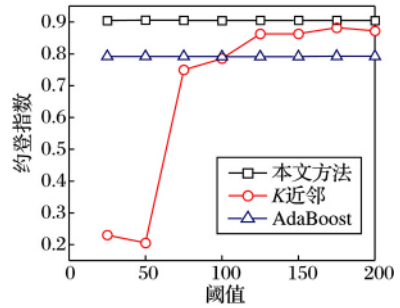

图 8 三种算法约登指数的比较

本文方法与 K 近邻和 AdaBoost 分类器的对比结果表明，本文算法在异常检测分类中表现最佳，性能稳定，在选择较低阈值时就能达到较高检测结果，与其他两种算法相比大大降低时间复杂度。三种算法的检测率都达到 90%，但其他两种分类器在误报率、阳性似然比和约登指数三类指标中的表现远不如本文方法。

综合以上实验结果，本文方法在检测率和检测时间上均优于随机森林分类器，其检测率较随机森林提高约 0.2%，训练过程和预测过程的时间缩短比率约为 20% 和 17%，时间效率明显提高。本文算法与 K 近邻和 AdaBoost 算法相比，三类算法在检测率上都能达到 90%，但本文方法在误报率、阳性似然比和约登指数三项指标上均优于其他算法，且本文算法更为稳定。

4 结语

针对高维数据对网络异常检测的检测率和检测时间产生不利影响的问题,提出一种基于信息增益特征选择的网络异常检测模型,使用信息增益算法对数据进行降维,提高随机森林分类器的检测率、缩短检测时间。实验结果表明,本文方法在检测率和时间效率上均优于随机森林分类器;在误报率、阳性似然比和约登指数上均比 K 近邻和 AdaBoost 算法效果更好。

本文使用现有数据集进行研究和分析,下一步的研究重点是采集现网数据进行进一步分析。

参考文献:

[1] 杨宏宇,朱丹,谢丰,等. 入侵异常检测研究综述[J]. 电子科技大学学报, 2009, 38(5): 587 – 596.

[2] 杨雅辉,黄海珍,沈晴霓,等. 基于增量式 GHSOM 神经网络模型的入侵检测研究[J]. 计算机学报, 2014, 37(5): 1216 – 1224.

[3] 穆祥昆,王劲松,薛羽丰,等. 基于活跃熵的网络异常流量检测方法[J]. 通信学报, 2013(S2): 51 – 57.

[4] SINGH R, KUMAR H, SINGLA R K. An intrusion detection system using network traffic profiling and online sequential extreme learning machine[J]. Expert Systems with Applications, 2015, 42(22): 8609 – 8624.

[5] KUANG F, ZHANG S, JIN Z, et al. A novel SVM by combining kernel principal component analysis and improved chaotic particle swarm optimization for intrusion detection[J]. Soft Computing, 2015, 19(5): 1187 – 1199.

[6] LI J, ZHANG S, LU Y, et al. Internet traffic classification using machine learning[C]// Proceedings of the 2007 Second International Conference on Communications and Networking in China. Piscataway: IEEE, 2007: 68 – 72.

[7] RONAN C A, CHO S B. Mining SQL queries to detect anomalous database access using random forest and PCA[C]// Current Approaches in Applied Artificial Intelligence, LNCS 910. Berlin: Springer-Verlag Berlin, 2015: 151 – 160.

[8] DU Y, ZHANG R, LIU J, et al. An optimized anomaly intrusion detection scheme using KNN algorithm[J]. Journal of Computational & Theoretical Nanoscience, 2011, 4(8): 2941 – 2945.

[9] ZENG Y, CHEN T M. Classification of traffic flows into QoS classes by unsupervised learning and KNN clustering[J]. Ksii Transactions on Internet and Information Systems, 2009, 3(2): 134 – 146.

[10] CAO L C. Detecting Web-based attacks by machine learning[C]// Proceedings of 2006 International Conference on Machine Learning and Cybernetics. Piscataway: IEEE, 2006: 2737 – 2742.

[11] 董超,周刚,刘玉娇,等. 基于改进的 Adaboost 算法在网络入侵检测中的应用[J]. 四川大学学报(自然科学版), 2015, 52(6): 1225 – 1229.

[12] HARRINGTON P. Machine Learning in Action[M]. Greenwith, CT, USA: Manning Publications, 2012.

[13] 何慧, 苏一丹, 覃华. 基于信息增益的贝叶斯入侵检测模型优化的研究[J]. 计算机工程与科学, 2006, 28(6): 38 – 40.

[14] SANTIAGOPAZ J, TORRESROMAN D, FIGUEROAYPIÑA A, et al. Using generalized entropies and OC-SVM with Mahalanobis kernel for detection and classification of anomalies in network traffic[J]. Entropy, 2015, 17(9): 6239 – 6257.

[15] TAN P-N, STEINBACH M, KUMAR V. 数据挖掘导论[M]. 范明,范宏建,等译. 北京: 人民邮电出版社, 2011.

[16] ZHANG J, ZULKERNINE M, HAQUE A. Random-forests-based network intrusion detection systems[J]. IEEE Transactions on Systems Man and Cybernetics Part C – Applications and Reviews, 2008, 38(5): 649 – 659.

[17] CHEN Y, WU W. A prospecting cost-benefit strategy for mineral potential mapping based on ROC curve analysis[J]. Ore Geology Reviews, 2016, 74: 26 – 38.

(上接第 40 页)

参考文献:

[1] 崔铁军, 马云东. 多维空间故障树构建及应用研究[J]. 中国安全科学学报, 2013, 23(4): 32 – 37.

[2] 章恩泽, 陈庆伟. 不确定可靠性优化问题的多目标粒子群优化算法[J]. 控制与决策, 2015, 30(9): 1169 – 1705.

[3] 高扬, 王义龙, 牟德一. 基于不确定理论和 CREAM 的飞行员应急操作可靠性分析[J]. 中国安全科学学报, 2013, 23(10): 56 – 62.

[4] 毛松, 师义民, 孙天宇. Pareto 产品可靠性实验最少试件数的确定[J]. 机械强度, 2013, 35(3): 274 – 277.

[5] 富立, 王新玲, 岳亚洲. 基于可靠性分析的最优冗余配置数量确定方法[J]. 北京航空航天大学学报, 2010, 36(9): 1030 – 1033.

[6] 许艾明, 赵柱, 陈琨, 等. 非确定工作状态下机械系统可靠性分析[J]. 机械设计与制造, 2012(1): 100 – 102.

[7] 崔铁军, 马云东. 空间故障树的径集域与割集域的定义与认识[J]. 中国安全科学学报, 2014, 24(4): 27 – 32.

[8] 崔铁军, 马云东. 基于多维空间事故树的维持系统可靠性方法研究[J]. 系统科学与数学, 2014, 34(6): 682 – 692.

[9] 崔铁军, 马云东. 宏观因素影响下的系统中元件重要性研究[J]. 数学的实践与认识, 2014, 44(18): 124 – 131.

[10] 崔铁军, 马云东. 考虑人因失误和状态检修的事故链式模型研究[J]. 中国安全科学学报, 2014, 24(8): 37 – 42.

[11] 崔铁军, 马云东. 考虑范围属性的系统安全分类决策规则研究[J]. 中国安全生产科学技术, 2014, 10(11): 6 – 9.

[12] 崔铁军, 马云东. 考虑点和线的有向无环网络连通可靠性研究[J]. 计算机应用研究, 2015, 32(11): 3315 – 3318.

[13] 崔铁军, 马云东. 基于不完全维修的可修系统平均故障次数研究[J]. 系统工程理论与实践, 2016, 36(1): 184 – 188.

[14] 崔铁军, 马云东. 系统可靠性决策规则发掘方法研究[J]. 系统工程理论与实践, 2015, 35(12): 3210 – 3216.

[15] 崔铁军, 马云东. 连续型空间故障树中因素重要度分布的定义与认知[J]. 中国安全科学学报, 2015, 25(3): 23 – 28.

[16] 李莎莎, 崔铁军, 马云东. 基于空间故障树理论的系统可靠性评估方法研究[J]. 中国安全生产科学技术, 2015, 11(6): 68 – 74.

[17] 李德毅, 杜鹢. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005.

[18] 李健, 汪明武, 徐鹏, 等. 基于云模型的围岩稳定性分类[J]. 岩土工程学报, 2014, 36(1): 83 – 87.

[19] 罗赟骞, 夏靖波, 陈天平. 基于云模型和熵权的网络性能综合评估模型[J]. 重庆邮电大学学报(自然科学版), 2009, 21(6): 771 – 775.

[20] 陈昊, 李兵. 基于逆向云和概念提升的定性评价方法[J]. 武汉大学学报(理学版), 2010, 56(6): 683 – 688.