doi:10.3969/j.issn.1673-9833.2014.03.014

基于贝叶斯方法的网络安全态势感知模型

叶健健，文志诚，吴欣欣

（湖南工业大学 计算机与通信学院，湖南 株洲 412007）

摘 要：提出了一种基于贝叶斯方法的网络安全态势感知模型。分析了目前国内外NSSA模型的研究现状。本模型先对历史监测数据进行分析，得到先验概率，随后以构建的时序模型为时间推进过程，利用贝叶斯方法对数据进行处理，将历史的统计数据与监测数据相结合，进行有效的安全预测。在网络结构的构建上，采用层次化结构，配合较合理的评价体系，使得该模型能够准确、快速、合理的对网络安全状态进行预测，并且具有较好的实时性。

关键词：贝叶斯分析方法；网络安全态势感知；时序分析

中图分类号：TP393.08 **文献标志码**：A **文章编号**：1673-9833(2014)03-0065-06

The Network Security Situational Awareness Model Based on Bayesian Method

Ye Jianjian，Wen Zhicheng，Wu Xinxin

（School of Computer and Communication，Hunan University of Technology，Zhuzhou Hunan 412007，China）

Abstract：Proposes a network security situational awareness model based on Bayesian method, and analyzes the research status of NSSA model at home and abroad. Analyzes historical monitoring data and obtains prior probability, and then constructs the time-series model for the process of time, applies Bayesian method to process the data, combines the historical statistical data with monitoring data and conducts the effective safety prediction. On the network structure construction, uses a hierarchical structure and coordinates a reasonable evaluation system for the model accurately, fast and reasonable prediction of network security, and the model has better real-time performance.

Keywords：Bayesian analysis; network security situational awareness; timing analysis

1 相关研究

网络的快速发展给人们带来了便利，但随之而来的网络安全问题也越来越突出。网络攻击普遍存在于各种环境中，并且日趋规模化、分布化、复杂化，破坏性也越来越强，还能造成巨大的经济损失。

传统的安全手段主要以被动式防御为主，一般是在发生网络安全事故后，才进行防御。这种方法已经很难满足新形势下的安全需求。在这种情况下，迫切需要一种新的安全技术来保障网络的安全。网络安全态势感知（network security situation awareness，NSSA）[1]能够实时地感知网络威胁和风险，使得安全分析员能够准确地感知网络安全状况，从而及时、准确地作出决策，将网络中不安定因素带来的风险和损失降到最低限度。针对不同的网络安全问题，国内外研究学者提出了多种网络安全态势评估方法。

SIFT（security incident fusion tool）项目组研制了Nvision IP[2]和Vis-Flow Connect[3]2种可视化工具。Nvision IP可以显示一个B类网络的连接状态，并且

收稿日期：2014-02-30

作者简介：叶健健（1989-），男，广东南雄人，湖南工业大学硕士生，主要研究方向为网络安全态势感知，
E-mail：464669778@qq.com

提供了3种不同精度的视图；Vis-Flow Connect能显示一个局域网络内、外各主机之间的通讯连接信息，该类工具仅反映了网络连接状态，评估指标较为单一，对管理员的经验水平要求较高。

T. Bass[4]提出了利用入侵检测系统（intrusion detection systems，IDS）的分布式传感器进行数据融合的方法，对计算机网络安全态势进行评估，通过数据融合和数据挖掘的方法评估计算机网络的安全性，但没有实现具体的原型系统。

Information Extraction & Transport开发了SSARE（security situation assessment and response evaluation）系统[5]，用于广域网的攻击检测、态势评估和响应评估。该系统实现了入侵检测、态势评估和响应评估的有机结合，但是信息获取方式较为单一。

V. Gorodetsky等人[6]提出了基于异步数据流的网络安全态势评估方法，利用多代理异常检测网络的数据流并进行分析，获取安全态势。但是，该方法只考虑攻击信息，而忽略了网络本身的特性。

V. Yegneswaran等人[7]提出了利用Honeynets进行安全态势评估的方法。该方法利用Honeynets提供的大量网络活动信息，根据入侵检测工具Bro对这些活动产生的报警信息来构建安全态势曲线，但只有在大规模病毒或蠕虫爆发时，该曲线才能体现出明显的效果。

陈秀真等人[8]提出了层次化网络安全威胁态势量化评估方法。该方法利用IDS的报警信息和网络性能指标，结合服务、主机本身的重要性及网络系统的组织结构，采用自下而上、先局部后整体的评估策略，将网络分为服务、主机、系统来进行分层计算；最后，通过综合分析，得到网络安全态势图，并有集成化的系统实现，其具有较好的理论和实用价值。

张海霞等人[9]提出了基于攻击能力增长的网络安全分析模型。该模型将攻击能力增长表示攻击者的最终目标，使得攻击图的表示更为准确，并以此分析攻击路径，从而进行网络安全性的分析。

王超等人[10]提出了一种基于隐马尔可夫模型（hidden markov model，HMM）的内部人员资源滥用行为检测方法。该方法以信息系统的敏感文件夹作为模型的状态，以用户的事务处理操作作为观测符号，采用Baum Welch算法确定模型参数，基于该模型建立内部人员访问行为的HMM模型，并用于资源滥用行为检测。

此外，还有其他学者从不同的思路建立了NSSA模型。有的从层次化模型[11]进行分析，模型逻辑清晰；有的对数据的融合[12-14]进行研究，具有独到的特色。

以上这些国外内学者为网络安全态势评估提供了有效的解决方法，为评估模型及算法的研究奠定了良好的基础。但是，这些方法都具有片面性，例如：缺乏对网络安全因素的全面考虑，评估数据源单一，使得评估结果缺乏全面性；忽略了数据源之间的互补性和冗余性等内在联系，使得评估结果不够准确；另外，这些方法无法对安全状况的发展趋势进行预测分析。

本文提出了一种基于贝叶斯方法的网络安全态势感知模型。该模型在时序模型的基础上，利用贝叶斯方法将历史的统计数据与检测数据相结合，进行有效地安全预测，具有较好的实时性。该方法不仅具有清晰的网络构造，而且可以对数据进行快速、准确地处理，对未知数据的预测可靠性更高。

2 NSSA模型概述

2.1 态势感知系统框架

态势感知（situation awareness）这一概念源于航天飞行的人因（human factors）研究[1]，此后，被广泛应用于军事战场、核反应控制、空中交通监管（air traffic control，ATC）以及医疗应急调度等领域。态势感知这项课题越来越被人重视，是因为在动态复杂的环境中，决策者需要借助态势感知的方法和工具，显示当前环境的连续变化状况，以便准确地作出决策。

1988年，M. R. Endsley在文献[15]中将态势感知定义为“在一定的时空条件下，对环境因素的获取、理解以及对未来状态的预测”。整个态势感知过程可由态势要素获取、态势理解和态势预测3级模型表示，如图1所示。

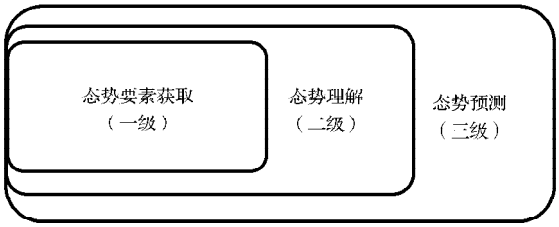

图1 态势感知过程图

Fig. 1 Diagram of situational awareness process

2.2 网络安全态势感知的相关技术

根据态势感知模型结构的3个层次的划分，在网络安全态势感知领域中也将可以借鉴。由此所产生的问题，可利用以下几种相关的技术来解决。

数据采集：数据采集属于底层部分，通过多传感器监测网络系统的运行状况，检测和收集大量的原

始安全数据。

态势理解：收集原始数据之后，需要对数据进行诠释，可采用规范化分析、冗余检测和冲突检测等方法，分析原始数据，得到规范化的数据集。

态势评估：采用态势评估算法，分析态势理解模块的数据，定量描述系统的安全态势，并构建合理的评价体系，对态势给出一个定量的评估。

态势预测：采用态势预测算法，分析态势的变化规律，预测系统安全态势变化趋势。

2.3　贝叶斯分析方法

贝叶斯分析方法（Bayesian analysis）提供了一种计算假设概率的方法，这种方法是基于假设的先验概率、给定假设下观察到不同数据的概率以及观察到的数据本身而得出的。该方法是，将关于未知参数的先验信息与样本信息综合，再根据贝叶斯公式，得出后验信息，然后根据后验信息去推断未知参数。贝叶斯公式为

$$P(C_j \mid X)=\frac{P(C_j)P(X \mid C_j)}{P(X)},$$

式中：$P(C_j \mid X)$为事件C_j在事件X发生的前提下的后验概率；$P(X \mid C_j)$为先验概率；$P(C_j)$为事件C_j发生的概率；$P(X)$为事件X发生的概率。

数据处理步骤如下：首先，统计历史数据，得到每种威胁发生的频率，将其作为先验概率；然后，利用贝叶斯公式将先验概率与样本信息综合，得到后验概率，其中，样本信息为实时监控得到的数据；以不同时段为分隔计算点，将得到的后验概率作为新一轮计算的先验概率，与进一步获得的样本信息综合，求得后验概率，其计算过程是迭代的。随着迭代过程的继续，后验概率将越来越复合实际情况。本文通过各种威胁发生的概率对网络安全进行评估，并对未来时刻的网络安全状态作出预测，以便更改相应的安全防护策略。

3　基于时序分析模型

时间序列分析技术是通过对预测目标自身时间序列的处理，来研究其变化趋势的方法。本文对历史的监测数据进行分析，将数据中各种攻击发生的时间作为时间序列，对一天的时间进行合理地分割，按照时间序列的时序建立模型。

3.1　模型的相关定义

网络攻击A：对引发 IDS 产生报警的黑客攻击行为表示为A={Name, Time, Type, SIP, DIP, SP, DP, Pro, Priority}，其中 Name, Time, Type 分别表示为攻击名称、发生时间以及攻击类型，SIP 和 DIP 代表源地址和目的地址；SP，DP 代表源端口和目的端口；Pro 表示协议类型；Priority 代表攻击威胁级别。

威胁指数R：攻击A对主机的安全威胁程度，攻击的威胁指数和攻击的成功与否、攻击带来的后果有直接关系。

主机的安全性H：每台主机存在多个不同的服务程序，不同的服务程序所受到的攻击类型及其带来的后果不尽相同。在给予不同的服务不同的权限下，所有攻击对主机的危害程度决定其安全性。

成本函数C：不同的服务所带来的效益不同，在受到攻击时，所产生的损失也不一样。攻击成功后，所有有效的攻击给主机造成的损失为成本。

网络安全系数L：网络中有各类不同功能的主机，他们有不同的权重系数。通过给定的安全阀值来判断网络状态，得到网络安全系数。该系数有助于网络管理员进行有效的策略制定，从而预防接下来发生的安全事故。

3.2　网络结构建模

网络的结构按照模型的相关定义进行层次化建模。该模型具有鲜明的层次结构，比较容易理解，并且符合实际的结构关系，自顶向下包括 Internet 网络、网络中的各个主机、每台主机中运行的各项服务，最底层为网络攻击，这些攻击针对不同的服务，如图 2 所示。

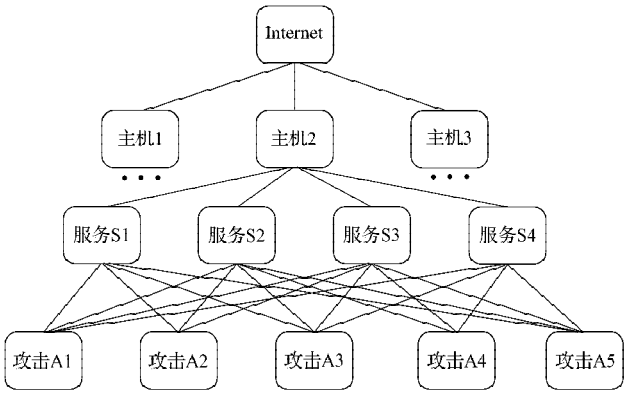

图 2　网络整体结构图

Fig. 2　The overall network structure

3.3　时序分析方法

本文将一天的时间划分为 24 等份，即时间粒度为一小时，分别为T_1, T_2,…, T_{24}。历史统计数据为当前历史状态P_p，如果下一刻主机收到攻击，则重新计算当前的安全系数P_n。例如：当前历史状态由指数 1 描述，实时监测的样本数据由样本 1 描述，由这 2 个数据可得下一刻的历史状态即指数 2。一般，攻击具有连续性，因此，一次攻击发生后，下一时刻

再次攻击的可能性比较大。故而，数据处理应保持连续性。具体的系数计算流程如图3所示。

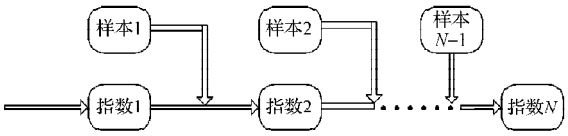

图3 系数计算流程

Fig. 3 Coefficient calculation process

3.4 量化分析公式

为方便分析网络的安全态势，必须对数据进行量化处理。构建评价体系[16]的具体实现如下。

1）攻击 A 的发生概率

$$P_{R_A}=\alpha P_{\mathrm{p}R_A}+(1-\alpha)N(t)\ , \qquad (1)$$

式中：$P_{\mathrm{p}R_A}$为攻击 A 在 t 时刻的发生概率；$N(t)$为当前监测样本发生次数；α为修正参数，将历史统计数据与当前数据结合。

2）攻击 A 的威胁指数 R

$$R_A(t)=f\left(\theta,N(t),B(t),D,P_{R_A}\right)=P_{R_A}\theta\left(N(t)10^D+100B(t)10^D\right)\ , \qquad (2)$$

式中：θ为正常访问量；$B(t)$为网络带宽占有率；D 为威胁的等级。

3）主机的安全性 H

$$\boldsymbol{H}(t)=f\left(\overrightarrow{\boldsymbol{R}_A}(t),\overrightarrow{\boldsymbol{W}}\right)=\overrightarrow{\boldsymbol{W}}\cdot\overrightarrow{\boldsymbol{R}_A}(t), \qquad (3)$$

式中，$\overrightarrow{\boldsymbol{R}_A}(t)=\left(R_{A_1},\cdots,R_{A_n}\right)$，$\overrightarrow{\boldsymbol{W}}$分别为 t 时刻攻击 A_i 的威胁指数和所对应服务的权重向量。

4）成本函数 C

$$C(t)=f\left(\overrightarrow{\boldsymbol{N}}(t),\overrightarrow{\boldsymbol{E}}\right)=\sum_{i=1}^{n}N_{A_i}(t)E_{A_i}\ , \qquad (4)$$

式中：$\overrightarrow{\boldsymbol{N}}(t)=\left(N_{A_i},\cdots,N_{A_n}\right)$，其中 N_{A_i} 为攻击 A_i 发生的次数；E_{A_i} 为攻击 A_i 所造成的损失。

5）网络安全系数 L

$$L(t)=f\left(\overrightarrow{\boldsymbol{H}}(t),\overrightarrow{\boldsymbol{C}}(t)\right)=\sum_{k=1}^{n}\left[\omega H_k(t)+(1-\omega)C_k(t)\right], \qquad (5)$$

式中，ω为 $\overrightarrow{\boldsymbol{H}}(t)$ 与 $\overrightarrow{\boldsymbol{C}}(t)$ 之间的权重参数。

4 测试结果与分析

本文以DARPA1999入侵检测数据集为基础，进行实验分析。以1999年3月1日—3月7日一个星期的监测数据为例，运用本算法对其进行安全威胁态势评估，分析这一个星期的各项主机服务、每台主机以及整个系统的安全威胁状态。

由于主机A中运行的主要是邮件服务和时间服务程序，故其它程序所受到的警报数量相对很少，本文只考虑这2种服务。以3月1日的数据为例，主机A发生警报按照24 h作图，如图4~5所示。由图4可知，在这一天中，邮件服务器发生警报的次数总体上保持平衡，都在100以内，只有在20~22 h时突发性地增加，由此可见，这个时间段是邮件服务器的危险高峰期。在图5中，时间在0~19 h，时间服务程序的警报量保持在一定的范围内，无太大波动，而19 h后，报警量有些增加。

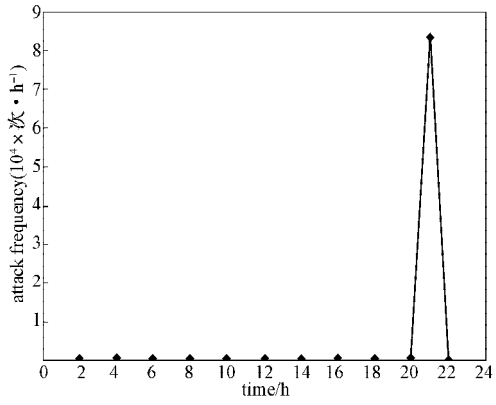

图4 邮件服务程序发生警报分布图

Fig. 4 The distribution of mail service procedure in alarm

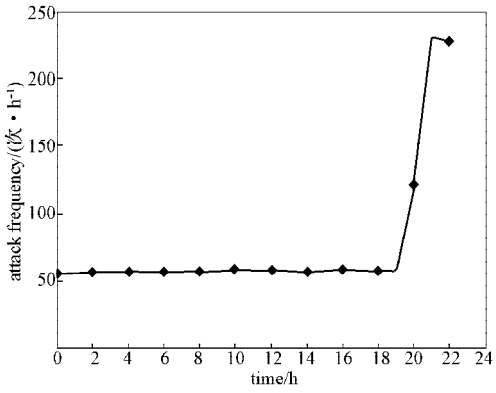

图5 时间服务程序发生警报分布图

Fig. 5 The distribution of time service procedure in alarm

图6为3月1日主机A的网络安全态势整体评价曲线。从图可以看出，由于邮件服务程序的安全态势波动较大，故而主机A整体受影响较大，在20~22 h这段时间的安全状况较差，可进行相应地防护。

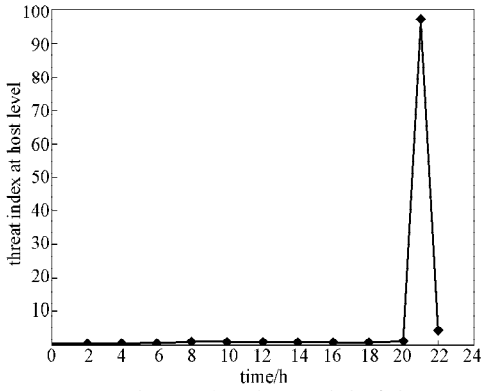

图6 主机A的安全威胁态势走势图

Fig. 6 The trend diagram for security threat of Host A

图7为整个星期的网络安全态势。由图可知，星

期一、星期六和星期天的攻击指数较高，而星期二～星期五的相对较缓，故而应该在周末前后加强网络安全的防护。

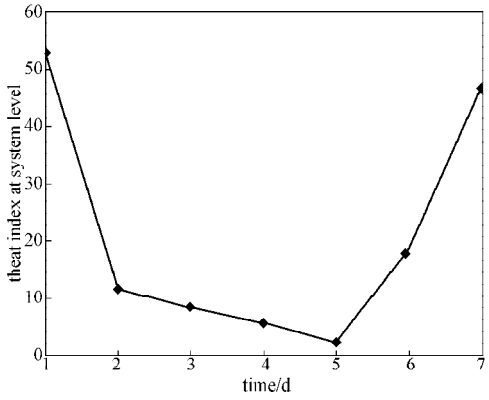

图7　网络安全态势评估图

Fig. 7　The assessment diagram for network security situation

通过攻击数据的分析，可以清楚地了解每一天、每一台主机所受到的攻击数量和安全评估指数，通过分析整周的安全态势评估，发现攻击的规律。从而根据集中发生攻击的时间段对主机加强管理，预防网络安全事故的发生。

在实际的监测应用中，可以根据实时的监测数据调整方案，以便适应动态的网络环境。

5　结语

本文提出了基于贝叶斯方法的网络安全态势感知模型。在结构的划分上，该模型采用清晰的层次化结构，其符合实际需求，且具有合理性；在数据的处理上，采用时序分析方法，按照时间的推进进行动态地调整，实时性较强，且利用贝叶斯方法，将历史的统计数据与检测数据相结合，进行有效地安全预测，具有快速、准确率高和实时性强的特点。

网络安全问题时刻威胁着我们，网络的变化也是越来越快，因此，本文收集的数据也是有限的。下一步的工作是，融合各种不同的监测数据进行全面的安全预测，且构建规模更大的网络结构。

参考文献：

[1] Theureau J. Use of Nuclear-Reactor Control Room Simulators in Research & Development[C]//7th International Federation of Automatic Control Symposium on Analysis，Design and Evaluation of Man-Machine Systems. Kyoto：[s. n.]，1998：425–430.

[2] Lakkaraju K，Yurcik W，Lee A J. NVisionIP：Netflow Visualizations of System State for Security Situational Awareness[C]//Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security. [S. l.]：ACM，2004：65–72.

[3] Yin X，Yurcik W，Treaster M，et al. VisFlowConnect：Netflow Visualizations of Link Relationships for Security Situational Awareness[C]//Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security. [S. l.]：ACM，2004：26–34.

[4] Bass T. Intrusion Detection Systems & Multisensor Data Fusion：Creating Cyberspace Situational Awareness[J]. Communications of the ACM，2000，43(4)：99–105.

[5] D′Ambrosio B，Takikawa M，Fitzgerald J，et al. Security Situation Assessment and Response Evaluation (SSARE) [C]//2001 DARPA Information Survivability Conference & Exposition II. Anaheim：IEEE，2001：387–394.

[6] Gorodetsky V，Karsaev O，Samoilov V. On-Line Update of Situation Assessment Based on Asynchronous Data Streams[C]//Knowledge-Based Intelligent Information and Engineering Systems. [S. l.]：Springer Berlin Heidelberg，2004：1136–1142.

[7] Yegneswaran V，Barford P，Paxson V. Using Honeynets for Internet Situational Awareness[C]// Proceedings of the Fourth Workshop on Hot Topics in Networks. Maryland：[s. n.]，2005：17–22.

[8] 陈秀真，郑庆华，管晓宏，等. 层次化网络安全威胁态势量化评估方法[J]. 软件学报，2006，17(4)：885–897.
Chen Xiuzhen，Zheng Qinghua，Guan Xiaohong，et al. Quantitative Hierarchical Threat Evaluation Model for Network Security[J]. Journal of Software，2006，17(4)：885–897.

[9] 张海霞，苏璞睿，冯登国. 基于攻击能力增长的网络安全分析模型[J]. 计算机研究与发展，2007，44(12)：2012–2019.
Zhang Haixia，Su Purui，Feng Dengguo. Network Security Analysis Model Based on the Increase in Attack Ability[J]. Journal of Computer Research and Development，2007，44(12)：2012–2019.

[10] 王　超，郭渊博，马建峰，等. 基于隐马尔可夫模型的资源滥用行为检测方法研究[J]. 电子学报，2010，38(6)：1383–1388.
Wang Chao，Guo Yuanbo，Ma Jianfeng，et al. HMM-Based Detection Method for Resource Misuse in Information Systems[J]. Acta Electronica Sinica，2010，38(6)：1383–1388.

[11] 郑黎明，邹　鹏，贾　焰. 多维多层次网络流量异常检测研究[J]. 计算机研究与发展，2011，48(8)：1506–1516.
Zheng Liming，Zou Peng，Jia Yan. Anomaly Detection Using Multi-Level and Multi-Dimensional Analyzing of Network Traffic[J]. Journal of Computer Research and Development，2011，48(8)：1506–1516.

[12] 韦　勇，连一峰. 基于日志审计与性能修正算法的网络

安全态势评估模型[J]. 计算机学报，2009，32(4)：763–772.
Wei Yong，Lian Yifeng. A Network Security Situational Awareness Model Based on Log Audit and Performance Correction[J]. Chinese Journal of Computers，2009，32(4)：763–772.

[13] 刘效武，王慧强，赖积保，等. 基于多源异质融合的网络安全态势生成与评价[J]. 系统仿真学报，2010，22(6)：1411–1415.
Liu Xiaowu，Wang Huiqiang，Lai Jibao，et al. Network Security Situation Generation and Evaluation Based on Heterogeneous Multi-Sensor Fusion[J]. Journal of System Simulation，2010，22(6)：1411–1415.

[14] 韦 勇，连一峰，冯登国. 基于信息融合的网络安全态势评估模型[J]. 计算机研究与发展，2009，46(3)：353–362.
Wei Yong，Lian Yifeng，Feng Dengguo. A Network Security Situational Awareness Model Based on Information Fusion[J]. Journal of Computer Research and Development，2009，46(3)：353–362.

[15] Endsley M R. Design and Evaluation for Situation Awareness Enhancement[C]//Proceedings of the Human Factors and Ergonomics Society Annual Meeting. [S. l.]：SAGE Publications，1988，32(2)：97–101.

[16] 王 娟，张凤荔，傅 翀，等. 网络态势感知中的指标体系研究[J]. 计算机应用，2007，27(8)：1907–1909.
Wang Juan，Zhang Fengli，Fu Chong，et al. Study on Index System in Network Situation Awareness[J]. Computer Applications，2007，27(8)：1907–1909.

[17] 王 嵘，刘 斌，刘东南，等. 动态网络分布式控制研究[J]. 湖南工业大学学报，2013，27(2)：63–67.
Wang Rong，Liu Bin，Liu Dongnan，et al. Research of Dynamic Distributed Control Networks，2013，27(2)：63–67.

（责任编辑：邓 彬）

（上接第23页）

Mainpulator for Humanoid Robots[J]. Robot，2011，33(3)：332–339.

[7] 成大先. 机械设计手册：第五卷[M]. 北京：机械工业出版社，2002：67–115.
Cheng Daxian. Mechanical Design Manual：Volume Fifth [M]. Beijing：Mechanical Industry Press，2002：67–115.

[8] Kurfess T R. Robotics and Automation Handbook[M]. Boca Raton：CRC Press，2005：112–322.

[9] 李大胜，缪鹏程. 遗传神经网络在数控机床刀具监测与控制系统中的应用[J]. 湖南工业大学学报，2013，27(3)：65–70.
Li Dasheng，Miao Pengcheng. Application of GA-BP Neural Network in CNC Machine Monitoring and Control System[J]. Journal of Hunan University of Technology，2013，27(3)：65–70.

（责任编辑：邓 彬）