

# 一种大数据条件下军事信息服务安全评估模型

张子春, 刘增良, 余达太

(北京科技大学 北京 100083)

**摘要:** 文中以大数据条件下的军事信息服务安全为评估对象,在粗糙集理论的基础上构建了合理的服务评价指标体系,建立了基于 BP 神经网络的服务安全评估模型,作为大数据条件下信息服务是否可行的评价依据,在整个服务的具体实施中指导军事信息服务的安全建设。

**关键词:** 大数据; 军事信息服务; 安全评估

**中图分类号:** G202      **文献标志码:** A      **文章编号:** 1009-8054(2014)06-0090-05

## A Security Evaluation Model for Military Information Service under Big-Data Condition

ZHANG Zi-chun, LIU Zeng-liang, YU Da-tai

(Beijing University of Science and Technology, Beijing 100083, China)

**Abstract:** With military information service as the evaluation object, a logical evaluation index system based on the rough set theory is built up, and then an evaluation model based on BP neural network technology also constructed. The model is regarded as the basis for evaluating the practicality of military information service under big-data condition, and guides the security construction of military information service during the whole process.

**Key words:** big data; military information service; security evaluation

## 0 引言

随着军事变革的不断深入以及信息技术的飞速发展, C/S 架构的指挥信息系统已不能满足联合作战的需求。尤其在栅格化信息网络基础设施的支撑下,下一代指挥信息系统的服务化发展成为必然。在此条件下,军事信息资源也遇到了前所未有的发展机遇,呈现出海量化、结构复杂化的大数据发展趋势<sup>[1]</sup>。大数据条件下的军事信息服务能够为各级指战员提供高效、有价值的信息服务,而安全问题是服务过程中一个尤为值得关注的要素,这也反映

了军事领域安全保密的特殊需求。文中从大数据条件下的军事信息服务安全性着眼,旨在建立一套科学、准确的安全评估方法,为大数据条件下军事信息服务是否可行提供安全性评价依据,并为其安全建设提供思路。

## 1 大数据条件下的军事信息服务安全特点

一般而言,军事信息服务属于 Web 服务范畴,具有 Web 服务的一般安全性特征。然而,大数据条件下的军事信息服务安全问题,由于其应用领域特点和数据环境特征,安全性与一般 Web 服务安全性有所差异,主要体现在两个方面。

### (1) 军事领域的安全保密性需求

1) 安全认证机制灵活。在栅格化信息网络环境下,军事信息服务的认证机制应更加灵活,能够支持不同的认证机制,支持单点登录,保证安全认证的及时与高效。

2) 保密策略相对独立。根据保密等级,军事信息服务

收稿日期: 2014-03-23

作者简介: 张子春, 1966 年生,男,博士研究生,研究方向为政治工作信息化; 刘增良, 1958 年生,男,教授,博士生导师,研究方向为信息对抗; 余达太, 1946 年生,男,博士生导师,研究方向为智能机器人。

应允许服务请求者与服务提供者制定各自独立的保密策略,并允许双方在安全策略协商的基础上建立安全通道。

3) 可靠性要求高。军事信息服务对软硬件设备的可靠性、网络传输的稳定性以及数据的真实性、实时性等因素要求更加严格。

## (2) 大数据带来的新型安全问题<sup>[2]</sup>

1) 用户隐私保护。大数据的特点之一大量收集用户信息进行行为预测,在为用户带来便利的同时也凸显了与用户隐私保护之间的冲突。军事信息服务涉及的用户隐私保护主要包括可控范围内的数据采集、数据共享以及销毁时的隐私保护问题。

2) 大数据可信性。大数据的可信性威胁主要来源于伪造数据或数据传输过程中的失真。对于军事信息服务来说,需要重点鉴别数据真实性、保证数据的完整性以及确保数据传输过程中的网络可靠性。

3) 大数据的访问控制。战场态势的动态变化导致用户角色和权限的频繁变更,角色和数据访问的授权机制更加动态灵活。

## 2 军事信息服务安全评估体系

### 2.1 安全评估指标

根据大数据条件下的军事信息服务的安全特点,文中以 WS - Security 标准<sup>[3]</sup>为基本框架,建立军事信息服务安全评估框架,根据安全特点调整相关指标,并在评估中增加相应权重。具体指标分类为:

1) 物理安全。物理安全包括介质安全性  $a_1$ 、环境安全性  $a_2$ 、设备安全性  $a_3$ 、设备可用性  $a_4$ 、设备可靠性  $a_5$ 、设备完整性  $a_6$ 、平台稳定性  $a_7$ 、抗电磁干扰指标  $a_8$ 、电气安全指标  $a_9$  等。

2) 网络安全。网络安全包括网络可用性  $a_{10}$ 、网络可靠性  $a_{11}$ 、网络完整性  $a_{12}$ 、网络机密性  $a_{13}$ 、网络真实性  $a_{14}$ 、抗抵赖性  $a_{15}$ 、传输可靠性  $a_{16}$  等。

3) 访问安全。访问安全包括身份验证  $a_{17}$ 、签名  $a_{18}$ 、加密  $a_{19}$  等。

4) 数据安全。数据安全包括数据完整性  $a_{20}$ 、数据可用性  $a_{21}$ 、数据真实性  $a_{22}$ 、数据实时性  $a_{23}$ 、数据私密性  $a_{24}$ 、用户隐私保护  $a_{25}$ 、服务信息  $a_{26}$  等。

5) 管理安全。管理安全<sup>[4]</sup>包括密钥管理  $a_{27}$ 、证书管理  $a_{28}$ 、人员安全意识  $a_{29}$ 。

对各安全性指标的测量,可根据 GJB142《军用软件安全性分析指南》及 MIG 平台相关标准进行安全等级划分,将等级通过隶属函数映射为数值,最终以 0 ~ 1 之间的数值表示,0 表示不安全,1 表示绝对安全,取值越高则安全程度越高。

### 2.2 基于粗糙集的指标约简

军事信息服务安全评估指标共列出了 5 类共 29 个指标,它们都在不同方面不同程度上反映了军事信息服务的安全程度。然而,在实际评价过程中,由于可能会存在评价因素过多、定性研究不足等问题,无法突出军事安全和大数据安全的需求特点,影响评价结果准确性和评价效率。为此,采用粗糙集理论(简称 RS)对影响军事信息服务安全的冗余因素进行约简,从而增加大数据条件下军事信息服务安全评估体系的合理性及科学性,也为后续 BP 神经网络模型的训练样本提供有效支持。

粗糙集理论(Rough Sets)是一种能够有效地处理不精确、不完整、模糊的信息并能发现隐含知识以及潜在规律的数学工具<sup>[5]</sup>,能够消除冗余数据,获得知识的最小表达,即实现决策表的属性约简。在粗糙集理论中,决策表被看作作为一种特殊的知识表达系统,同时具有条件属性和决策属性,形式化定义为:

#### 定义 1 决策表

设  $S = (U, A, V, f)$  是一种知识表达系统,论域  $U$  表示对象的有限非空集合,属性集  $A$  代表属性  $\alpha$  的有限非空集合,  $V$  为  $\alpha$  的值域,  $f$  为信息函数,其中  $A$  可由条件属性集  $C$  和决策属性集  $D$  组成,  $A = C \cup D$  且  $C \cap D = \emptyset$ ,那么这种具有条件属性  $\alpha$  ( $\alpha \in C$ ) 和决策属性  $d$  ( $d \in D$ ) 的知识表达系统则可被称为决策表或简称为 CD 决策表。

决策表的属性约简涉及到不可辨识关系的概念,它本质上是一种等价关系,形式化定义如下:

#### 定义 2 不可辨识关系

设  $S = (U, A, V, f)$  是一种知识表达系统,则  $\forall B \subseteq A$ ,其对应的不可辨识关系  $Ind(B) = \{ (x_i, x_j) \mid (x_i, x_j) \in U \times U, \forall b \in B (b(x_i) = b(x_j)) \}$ ,表示对象  $x_i$  和  $x_j$  关于属性集  $A$  的子集  $B$  是不可辨识的。

$Ind(B)$  的等价类称为属性子集  $B$  的基本集。

#### 定义 3 属性的独立性

若  $Ind(A) = Ind(A - \{a_i\})$ ,则说明属性  $a_i$  相对于属性集  $A$  而言是冗余的,否则认为属性  $a_i$  在  $A$  中是独立的,

即如果  $\forall a_i \in A, \text{Ind}(A - \{a_i\}) \neq \text{Ind}(A)$ , 则  $A$  是独立的。

定义 4 属性约简

设  $S = (U, A, V, f)$  是一种知识表达系统,  $\forall B \subseteq A$ , 如果  $B$  是独立的, 且  $\text{Ind}(B) = \text{Ind}(A)$ , 则称  $B$  是  $A$  的一个约简。

定义 5 属性的核

设  $S = (U, A, V, f)$  是一种知识表达系统, 属性集  $A$  的所有必要属性构成的集合记作  $\text{core}(A)$ 。记  $\text{red}(A)$  为  $A$  的所有约简集合, 那么存在关系  $\text{core}(A) = \bigcap \text{red}(A)$ , 即  $A$  的所有约简的交集构成  $A$  的核。

可见, 约简表示的是决策表的本质部分, 即它与原知识表达系统具有相同的辨识性, 能够辨别出所有原系统可辨别的对象。核则是所有约简的共同部分, 是进行约简时不能删除的部分。

定义 6 分辨矩阵

设  $S = (U, A, V, f)$  是一种知识表达系统,  $U = \{x_1, x_2, \dots, x_n\}$  为论域,  $A = C \cup D$  是属性集合, 子集  $C = \{a_i \mid i = 1, 2, \dots, m\}$  是条件属性集,  $D = \{d\}$  是决策属性集,  $a_k(x_j)$  是样本  $x_j$  在属性  $a_k$  上的取值。定义  $S$  的分辨矩阵为  $M(S) = [m_{ij}]_{n \times n}$ , 其  $i$  行  $j$  列的元素为:

$$m_{ij} = \begin{cases} a_k \in C, & a_k(x_i) \neq a_k(x_j) \wedge D(x_i) \neq D(x_j) \\ \phi, & D(x_i) = D(x_j) \end{cases} \quad i, j = 1, 2, \dots, n$$

不难看出, 分辨矩阵是一个对称矩阵, 在进行运算时, 只需计算其上三角或下三角部分即可。

定义 7 分辨函数

决策表  $S$  的分辨函数是一个具有  $m$  元变量  $a_1, a_2, \dots, a_m$  ( $a_i \in C, i = 1, 2, \dots, m$ ) 的布尔合取, 形式化表示为:  $f_{M(S)}(a_1, a_2, \dots, a_m) = \bigwedge \{ \bigvee m_{ij} \mid 1 \leq j < i \leq n, m_{ij} \neq \phi \}$ 。其中,  $\bigvee m_{ij}$  是矩阵项  $m_{ij}$  中各元素的析取, 分辨函数析取范式中每个析取项就对应一个约简。矩阵中所有单个元素构成的集合, 即  $\text{core}(A) = \{a_k \in A \mid m_{ij} = \{a_k\}, 1 \leq j < i \leq n\}$  则对应属性的核。

计算决策表  $S$  的约简算法为:

- 1) 计算  $S$  的分辨矩阵  $M(S)$ ;
- 2) 计算  $M(S)$  所对应的分辨函数  $f_{M(S)}$ ;
- 3) 计算  $f_{M(S)}$  的最小析取范式, 其中每个析取分量对应一个约简。

显然, 对军事信息服务安全评估指标进行属性约简, 就是约简评估指标构成的条件属性, 使其获得与原指标集具有相同辨识性的最小集, 简而言之, 就是使约简后的指

标体系能够以更少的条件属性获得与约简前的指标体系相同的功能。

在此理论基础上, 对军事信息服务质量评价指标体系给出的指标进行简化, 筛选出重要程度较高的评价因子, 可以通过如下具体步骤完成:

1) 整理数据。表 1 中, 指标值的获取根据实际情况通过系统日志读取、监控软件读取以及人工评分 3 种方式。与系统运行直接相关的指标, 如服务信息, 可采用系统日志读取的结果; 可利用软件监控的指标, 如数据实时性, 采用监控软件提供的采集数据; 对于不能自动获取的指标, 比如人员安全意识, 则通过人工评分的方式获得。在初始状态下, 定量指标可直接以具体数值表示, 定性指标则可根据排序等级的方式进行量化处理后确定指标值。

表 1 军事信息服务安全评估指标体系

指标体系	一级指标	二级指标
军事 信息 服务 安全 评估 指标 体系	物理安全	环境安全性
		设备可用性
		设备可靠性
		设备完整性
		平台稳定性
	网络安全	抗电磁干扰指标
		网络可用性
		网络可靠性
		网络完整性
		网络机密性
	访问安全	抗抵赖性
		传输可靠性
		身份验证
		签名
		加密
	数据安全	数据完整性
		数据可用性
		数据真实性
		数据实时性
用户隐私保护		
管理安全	服务信息	
	密钥管理	
	证书管理	
		人员安全意识

2) 指标值离散归一化处理。由于决策表只能识别离散化数据, 因此需要对决策表中的各指标值进行离散化、归一化处理。需要把握的原则是: 在对指标值进行离散归

一化处理,属性值的种类要尽量少,信息丢失也要尽量少。对于连续型数据的处理,可以用字母或数字代替,比如,可将定量指标值划分等级,比如将评价结果划分为A、B、C、D、E这5个评价等级,其中,取值[0.9,1]为A;取值[0.8,0.89]为B;取值[0.7,0.79]为C;取值[0.6,0.69]为D;取值0.59以下则为E。

3) 建立决策表。将评价指标集 $\{a_1, a_2, a_3, \dots, a_{29}\}$ 作为条件属性集C,取n个专家的综合评分集作为决策属性集D,离散处理后的指标值 $a_i(x_j)$ 作为单元取值,建立二维决策表(见表2)。

表2 二维决策表

U	A				
	条件属性集				决策属性
	$a_1$	$a_2$	$\dots$	$a_{29}$	d
$x_1$	$a_1(x_1)$	$a_2(x_1)$	$\dots$	$a_{29}(x_1)$	$d_1$
$x_2$	$a_1(x_2)$	$a_2(x_2)$	$\dots$	$a_{29}(x_2)$	$d_2$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$x_n$	$a_1(x_n)$	$a_2(x_n)$	$\dots$	$a_{29}(x_n)$	$d_n$

4) 计算得出约简结果。根据前述的约简算法,构建分辨矩阵,计算分辨函数的最小析取式,得出约简后的指标集。

### 2.3 指标体系构建

根据某指挥信息系统的仿真数据,建立军事信息服务安全评估指标体系,如表1所示。

## 3 基于BP神经网络的评估模型

### 3.1 BP神经网络

BP神经网络(Back Propagation Neural Network)模型是一种多层前馈型人工神经网络,通过反向传播的方式进行权值调整<sup>[6]</sup>。

BP神经网络的学习过程为有导师监督的学习过程,整个过程由信号的正向传播过程和权值调整的反向传播过程共同构成<sup>[7]</sup>。首先需要预先准备一组预期状态的输入输出数据对,然后将输入样本加载到输入层中,经过正向传播由隐含层传到输出层,将实际输出结果与期望输出结果比较,若比较结果不符合要求,则转入误差的反向传播过程,即将二者的误差通过隐含层反向传到输入层,在此过程中,调整修正各连接权重和阈值<sup>[8]</sup>。随着输入样本的增多,神经网络不断朝着正确响应的方向演变,直至实际输出与预期输出之间的误差能够控制在允许的范围之

内。BP神经网络的拓扑结构由输入层、隐含层和输出层共同组成,其中隐含层的数量可以是一层或多层<sup>[9]</sup>,如图1所示。

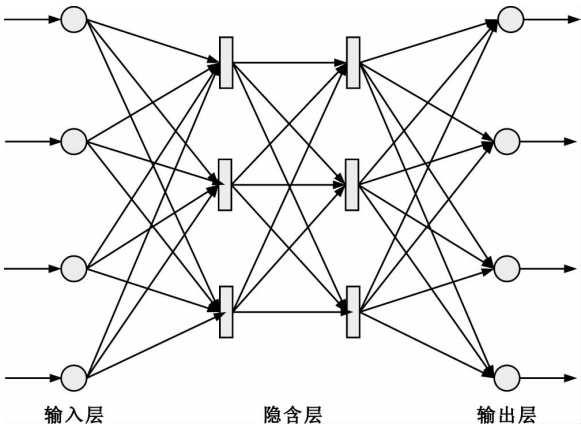


图1 BP神经网络拓扑结构

增加隐含层的层数,能够从理论上进一步减少误差,提高精度,但同时也会带来网络复杂度增加的问题<sup>[10]</sup>。在实际应用中,通过调整隐含层节点数目来提高精度比增加隐含层层数更加容易实现。总的来说,3层BP神经网络结构较易构建,并能方便地实现不同维度的近似映射,能够满足军事信息服务质量评价时效性强、准确度高的要求,因此,文中将采用3层BP神经网络拓扑结构,即仅设置一层隐含层。

### 3.2 模型设计

在此基础上,根据以下4个主要步骤构建基于BP神经网络的评价模型并实施评价过程。

#### (1) 设计输入层、输出层

根据上节步骤,用粗糙集对指标集进行预处理,得到的约简后的指标集作为BP神经网络输入层的节点,因此,输入层节点数应为约简后的指标个数。对于输出层而言,由于军事信息服务安全评估的最终结果只有一个,所以应该只有一个输出层节点。从数据库中选择一定数量的专家对各指标的评分作为训练样本输入,以每位专家的综合评分作为期望输出。

#### (2) 确定隐含层神经元个数

隐含层神经元个数的确定一般需要参考输入层与输出层神经元个数。普遍应用的计算公式是Kolmogorov定理给出的函数<sup>[7]</sup>:  $k = \sqrt{n+m} + b$ 。其中k表示隐含层神经元个数,n和m分别对应输入层和输出层神经元的个



数  $b$  为修正常量,取值范围为  $[1,10]$ 。经过计算,确定隐含层节点数的范围。在此基础上,需要采用试凑法进一步确定隐含层神经元个数,即通过对该范围内的节点数依次试验,最终得到隐含层神经元的最佳数量。由此,建立军事信息服务安全评估的神经网络模型,如图 2 所示。

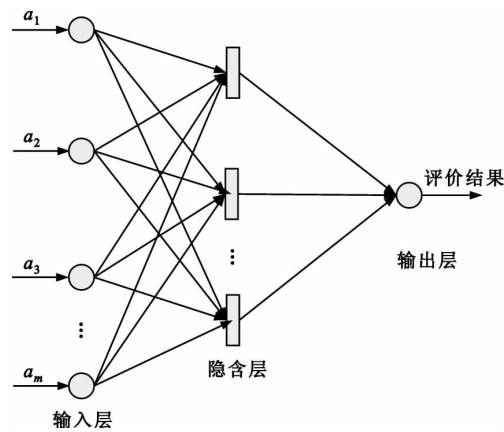


图2 军事信息服务安全评估 BP 神经网络模型

(3) 选择各层的激励函数

隐含层激励函数采用  $(0,1)$  S 型正切函数,如下所示:

$$f(x) = \frac{1}{1 + e^{-x}}$$
 输出层激励函数采用 Sigmoid 型对数函数。

(4) 选择误差函数

采用如下公式计算样本误差:  $E_p = \frac{1}{2} \sum_i (t_{pi} - O_{pi})^2$ 。

其中  $E_p$  表示第  $p$  个样本的误差值,  $t_{pi}$  表示期望输出值,  $O_{pi}$  表示计算输出值。

3.3 模型验证

根据某指挥信息系统的仿真数据,计算出决策表  $S$ , 确定模型输入层节点数为 24,隐含层节点数为 8,表 3 列出了部分训练样本数据。

通过 Matlab 神经网络工具箱计算,设定学习速率为 0.1,最大迭代次数 5 000,误差小于 0.001,经过 1 948 次训练,达到预定的误差范围 0.001,完成模型训练。利用训练好的模型对新输入的仿真数据进行计算,实际结果与期望结果能够达到基本吻合。

表3 军事信息服务安全评估模型部分训练样本数据

环境 安全性	设备 可靠性	网络 可用性	传输 可靠性	身份 验证	数据 完整性	数据 可用性	用户 隐私保护	密钥 管理	...	总评
0.8	0.83	0.91	0.84	1.0	0.89	0.95	0.75	0.93	...	A
0.91	0.94	0.89	0.85	0.94	0.89	0.87	0.78	0.81	...	B
0.74	0.92	0.91	0.92	0.87	0.95	0.96	0.85	0.89	...	B
0.62	0.84	0.51	0.56	0.88	0.74	0.71	0.59	0.65	...	D
0.58	0.93	0.76	0.87	0.98	0.86	0.75	0.67	0.92	...	B
0.78	0.93	0.92	0.89	0.87	0.75	0.82	0.78	0.62	...	B
0.88	0.86	0.83	0.77	0.93	0.94	0.91	0.59	0.80	...	A
0.93	0.88	0.85	0.75	0.91	0.87	0.81	0.87	0.88	...	B
0.69	0.54	0.87	0.64	0.76	0.96	0.89	0.96	0.92	...	B
0.85	0.91	0.89	0.88	0.85	0.81	0.9	0.81	0.95	...	A
0.93	0.94	0.89	0.88	0.91	0.75	0.92	0.75	0.97	...	B
0.91	0.93	0.93	0.76	0.85	0.82	0.62	0.82	0.93	...	A
0.87	0.81	0.86	0.88	0.85	0.91	0.80	0.95	0.92	...	B
0.94	0.89	0.92	0.69	0.78	0.87	0.88	0.76	0.91	...	C
0.88	0.9	0.95	0.87	0.89	0.96	0.92	0.87	0.94	...	A
0.91	0.92	0.92	0.96	0.95	0.81	0.76	0.89	0.87	...	A
0.59	0.62	0.71	0.81	0.94	0.75	0.87	0.77	0.96	...	C
0.78	0.80	0.82	0.75	0.86	0.82	0.89	0.75	0.91	...	B
0.81	0.88	0.92	0.82	0.75	0.62	0.77	0.64	0.75	...	B
0.96	0.92	0.93	0.91	0.94	0.93	0.75	0.88	0.82	...	A

(下转第 99 页)

的细节,只是一种地形网格细分。因此,下一步将对增加真实细节和实时的地形数据更新作进一步的研究。

#### 参考文献:

- [1] Duchaineau Mark ,Wolinsky Murray ,et al. ROAMing Terrain: Real - time Optimally Adapting Meshes [C]. In: Yagel R ,Hagen H ,eds. IEEE Visualization97. Los Alamitos: IEEE Press ,1997: 81 - 82.
- [2] Strugar F. Continuous Distance - dependent Level of Detail for Rendering Height Maps [J]. Journal of Graphics GPU and Game Tools 2009.
- [3] Jie Jiang ,Ling - Da Wu ,Rui Cao ,Bing Yang. Management and Web Publication of Global Satellite Remote - Sensing Images [C]. 7<sup>th</sup> International Conference on Information Communications and Signal Processing ( ICICS 2009) 2009.
- [4] Hakran Kim ,Yongik Yoon. Adaptation Method for Level of Detail ( LOD) of 3D Contents [J]. Network and Parallel Computing Workshops 2007.
- [5] 宋省身 ,全吉成 ,赵秀影 ,等. 基于 Chunked LOD 的实时细节增强地形算法 [J]. 计算机工程与设计 , 2014( 2) : 578 - 582.
- [6] SUN Yingying ,XU Miao. Research on Graphic Rendering of Grid Data [J]. Computer - Aided Industrial Design and Conceptual Design 2006.
- [7] LEVENBERG J. Fast View - dependent Level - of - detail Rendering using Cached Geometry [C]. VIS'02: Proceedings of the Conference on Visualization'02 2002. ■

(上接第 94 页)

## 4 结语

根据大数据条件下军事信息服务的特点,采用安全评估的方法指导军事信息服务的建设。采用粗糙集理论对安全指标集进行筛选简化,并在模拟数据的基础上确定了具有与原指标集相同辨识性的最小指标集。通过建立军事信息服务安全评估的 BP 神经网络模型,以约简后的指标集作为模型的输入样本,经过 BP 神经网络模型的训练确定指标权重,进行模型评价。实践证明,这是一个行之有效的方法。

#### 参考文献:

- [1] 城田真琴. 大数据的冲击 [M]. 周自恒,译. 北京: 人民邮电出版社 2013.
- [2] 冯登国,张敏,李昊. 大数据安全与隐私保护 [J]. 计算机学报 2014, 37( 1) : 246 - 258.
- [3] OASIS Web Services Security ( WSS) TC [EB/OL]. ( 2014 - 1 - 15) . [https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=wss](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wss).
- [4] 张磊,向德全,胥杰. 军用信息系统安全效能灰色评估模型和算法 [J]. 空军工程大学学报( 自然科学版) 2007, 8( 1) : 77 - 80.
- [5] 黄洪,余达太,刘增良. 粗糙集理论在规则抽取中的应用 [J]. 计算机工程与应用 2009, 45( 23) : 18 - 20.
- [6] 刘增良. 模糊技术与神经网络技术 [M]. 北京: 北京航空航天大学出版社 2008.
- [7] HU Guang - Da. Stability Criteria of Linear Neutral Systems with Distributed Delays [J]. Kybernetika , 2011, 47( 2) : 273 - 284.
- [8] 黄洪,刘增良,余达太. 一种智能化的数据分类、分级及保护模型 [J]. 北京工业大学学报, 2011, 37( 6) : 921 - 927.
- [9] 陶宇. 基于 BP 神经网络的 Web 服务评价模型研究 ( 硕士学位论文) [D]. 合肥: 安徽大学 2012.
- [10] Mohsen Saemi ,Morteza Ahmadi ,Ali Yazdian Varjani. Design of Neural Networks Using Genetic Algorithm for the Permeability Estimation of the Reservoir [J]. Journal of Petroleum Science and Engineering ,2007, 59( 1) : 97 - 105. ■