

## Elman 与 BP 神经网络在矿井水源判别中的应用

钱家忠, 吕 纯, 赵卫东, 潘 婧

(合肥工业大学 资源与环境工程学院, 合肥 230009)

**摘 要** 介绍了 Elman 神经网络与 BP 神经网络, 以谢一煤矿为例, 分别利用 Elman 网络与 BP 网络, 针对地下水化学特征分别建立突水判别模型, 实例结果表明: Elman 网络模型比 BP 网络模型具有更高的判别精度, 更快的运算速度, 更好的反应地下水系统特性, 为矿井水害防治提供了一种辅助决策手段.

**关键词** 突水水源; Elman 神经网络; BP 神经网络; 判别模型

## Comparison of application on Elman and BP neural networks in discriminating water bursting source of coal mine

QIAN Jia-zhong, LÜ Chun, ZHAO Wei-dong, PAN Jing

(School of Resources and Environmental Engineering, Hefei University of Technology, Hefei 230009, China)

**Abstract** The discrimination of the mine water-bursting source is deemed to be a basic knowledge for the water control in the mines. A speedy and precise discrimination is of key importance to the safe production of the whole mine. This paper introduces Elman neural networks and Back-propagation neural networks. Take Xieyi mine as an example, establishes the distinguishing model for water bursting by Elman neural networks and BP neural networks with groundwater chemical characteristics, respectively. Experimental results show that the Elman neural model is more precise and faster than BP neural model in discrimination. The Elman neural model could better respond characteristics of groundwater systems. It provides an assistant means for decision-making to prevent water-inrush from coal floor.

**Keywords** water bursting source; Elman neural networks; back-propagation neural networks; discriminating model

### 1 引言

矿井突水是煤矿生产过程最具威胁的自然灾害之一, 一旦矿井发生突水, 如何及时准确地判断突水成因, 查找突水水源, 是解决和进一步预防突水灾害的关键问题<sup>[1]</sup>. 因此, 准确判别矿井水源是矿井防治水的前提. 水化学成分能够反映地下水运移最本质的特征, 用水质资料判别水源具有快速、准确、经济的特点, 进行水源判别时, 由于各个含水层特性界限不明显, 具有很大的模糊性, 不确定性, 很难根据单个因子进行判别, 需要综合多个因素进行判别. 人工神经网络是对人脑或自然的神经网络若干基本特性的抽象和模拟, 是一种非线性的动力学系统. 它有大规模的并行处理和分布式的信息存储能力, 良好的自适应性、自组织性和很强的学习、联想、容错及抗干扰能力, 在判别这类具有模糊性的问题上有明显的优势<sup>[2]</sup>. 在人工神经网络中, BP 神经网络一般应用较多, 然而越来越多研究表明 BP 神经网络学习速度很慢, 失败性可能较大, 而且是一种常规静态模型<sup>[3]</sup>. 煤矿的地下水在几十年里, 各含水层的水质是有一定变化的, 使用一个静态的模型有些不合适, 而 Elman 神经网络是动态回归网络, 具有良好的动态模拟作用. 目前国内使用神经网络来进行水源判别的模型都是 BP 神经网络模型.

**收稿日期:** 2008-10-16

**资助项目:** 国家自然科学基金 (40672154); 教育部新世纪优秀人才支持计划 (NCET-06-0541)

**作者简介:** 钱家忠 (1968-), 教授, 博士生导师, 主要从事地下水模拟研究, E-mail: qjiazhong@gmail.com.

## 2 神经网络模型

神经网络是由大量的处理单元(神经元)互相链接而成的网络. 为了模拟大脑的基本特性, 在神经科学研究的基础上, 提出了神经网络模型. 但是, 实际上神经网络并没有完全反映大脑的功能, 只是对生物神经网络进行某种抽象、简化和模拟. 神经网络的信息处理通过神经元的相互作用来实现, 只是与信息的存储表现为网络原件互连分布式的物理联系. 神经网络的学习和识别取决于各神经元连接权系数的动态演化过程. 神经网络模型中主要有前向型神经网络和反馈型神经网络, 前向型神经网络中最常见的就是 BP 神经网络; 反馈型神经网络常见的有 Elman 神经网络和 Hopfield 神经网络.

### 2.1 Elman 神经网络

Elman 神经网络是 Elman 于 1990 年提出的, 该模型在前馈网络的隐含层中增加一个承接层, 作为一步演示算子, 达到记忆的目的, 从而使系统具有适应时变特性的能力, 能直接反应动态过程系统的特性.

#### 2.1.1 Elman 神经网络结构

Elman 型回归神经网络一般分为 4 层: 输入层、中间层(隐含层)、承接层和输出层, 如图 1 所示. 其输入层、隐含层和输出层的连接类似于前馈网络, 输入层的单元仅起到信号传输作用, 输出层单元起线性加权作用. 隐含层单元的传递函数可采用线性或非线性函数, 承接层又称为上下文层或状态层, 它用来记忆隐含层单元前一时刻的输出值, 可以认为是一个一步延时算子<sup>[4]</sup>. Elman 网络结构见图 1.

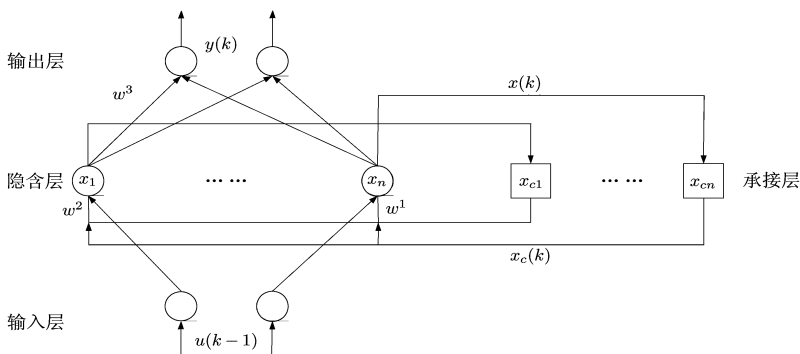


图 1 Elman 网络结构示意图

Elman 型回归神经网络的特点是隐含层的输出通过承接层的延迟与储存, 自联到隐含层的输入, 这种自联方式使其对历史状态的数据具有敏感性, 内部反馈网络的加入增加了网络本身处理动态信息的能力, 从而达到了动态建模的目的<sup>[5]</sup>. 利用煤矿几十年积累的数据, 按时间顺序对数据进行建模, 承接层记忆隐含层前一时刻的输出值, 对于煤矿地下水无规律的变化是十分有利的, 减少了最终输出的误差.

#### 2.1.2 Elman 神经网络的学习过程

以图 1 为例, Elman 网络的非线性状态空间表达式为:

$$y(k) = g(w^3 x(k)), \quad x(k) = f(w^1 x_c(k) + w^2(u(k-1))), \quad x_c(k) = x(k-1).$$

其中,  $y, x, u, x_c$  分别表示  $m$  维输出结点向量,  $n$  维中间层结点单元向量,  $r$  维输入向量和  $n$  维反馈状态向量.  $w^3, w^2, w^1$  分别表示中间层到输出层、输入层到中间层、承接层到中间层的连接权值.  $g(\cdot)$  为输出神经元的传递函数, 是中间层输出的线性组合.  $f(\cdot)$  为中间层神经元的传递函数, 常采用  $S$  函数.

Elman 网络也采用 BP 算法进行修正, 学习指标函数采用误差平方和函数

$$E(w) = \sum_{k=1}^n [y_k(w) - \tilde{y}_k(w)]^2,$$

其中  $\tilde{y}_k(w)$  为目标输出量.

Elman 神经网络学习算法采用的是优化的梯度下降算法, 即自适应学习速率动量梯度下降反向传播算法, 它既能提高网络的训练效率, 又能有效抑制网络陷入局部极小点. 学习的目的是用网络的实际输出值与输出样本值的差值来修改权值和阈值, 使得网络输出层的误差平方和最小.

### 2.2 BP 神经网络

人工神经网络 (ANNs) 被广泛用于各种领域, 20 世纪 80 年代中期, David Rumelhart, Geoffrey Hinton 和 Ronald Williams, David Parker, 以及 Yann Le Cun 分别独立地发现了 BP 算法.

### 2.2.1 BP 网络结构

BP 网络是一种单向传播的多层前向网络, 其结构如图 2 所示. 由图可见, BP 网络是一种具有三层或三层以上的神经网络, 包括输入层、中间层(隐层)和输出层. 上下层之间实现全连接, 而每层神经元之间无连接. 当一对学习样本提供给网络后, 神经元的激活值从输入层经各中间层向输出层传播, 在输出层的各神经元获得网络的输入响应. 接下来, 按照减少目标输出与实际误差的方向, 从输出层经过各中间层逐层修正各连接权值, 最后回到输入层, 这种算法称为“误差逆传播法”, 即 BP 算法. 随着这种误差逆的传播修正不断进行, 网络对输入模式响应的正确率也不断上升.

与感知器模型不同的是, BP 网络的传递函数要求必须是可微的, 所以不能使用感知器网络中的二值函数, 常用的有 Sigmoid 型的对数、正切函数或者线性函数. 由于传递函数是处处可微的, 所以对于 BP 网络来说, 一方面, 所划分的区域不再是一个线性划分, 而是由一个非线性超平面组成的区域, 它比较平滑的曲面, 因而它的分类比线性划分更加精确, 容错性也比线性划分更好; 另一方面, 网络可以严格采用梯度下降法进行学习, 权值修正的解析式十分明确.

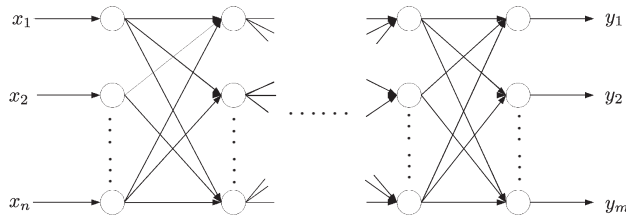


图 2 BP 网络结构示意图

### 2.2.2 BP 网络学习

神经网络的训练过程实际上是由一系列的步骤来完成的. 在学习阶段, 连接权重是由输入输出值来调整的<sup>[6]</sup>.

BP 神经网络是通过输出层的误差反向传播到隐含层来获得输出量的期望值. 梯度下降法利用计算网络权重和调整连接权重来最小化输出误差的. 神经网络误差函数的定义为

$$E = \frac{1}{2} \sum (T_k - O_k)^2,$$

$T_k$  和  $O_k$  分别是目标值和输出值.

梯度下降法则通过梯度误差改变权重

$$\Delta W_{ij} = -\eta \times \frac{\partial E}{\partial W_{ij}},$$

其中  $\eta$  是学习速率,  $\partial E / \partial W_{ij}$  由下列表示:

$$\frac{\partial E}{\partial W_{ij}} = -\delta_j^n \cdot A_i^{n-1}.$$

由上式我们获得梯度误差为

$$\Delta W_{ij} = -\eta \cdot \delta_j^n \cdot A_i^{n-1}.$$

这里, 输出层  $A_i^{n-1}$  与连接权重 ( $W_{ij}$ ) 有关,  $\delta_j^n$  是误差信号, 它可以通过计算来判断  $j$  是不是输出层的神经元. 如果  $j$  是输出层的神经元, 那么

$$\delta_j = (T_j - Y_j) \cdot Y_j \cdot (1 - Y_j).$$

如果  $j$  是隐含层的神经元, 则

$$\delta_j = \left[ \sum_j \delta_j \cdot (W_{hly})_{hj} \right] \cdot H_h \cdot (1 - H_h),$$

这里  $H_h$  是隐含层的数值. 最后, 连接神经元的权重值可以表达为<sup>[7]</sup>

$$W_{ij}^m = W_{ij}^{m-1} + \Delta W_{ij}^m = W_{ij}^{m-1} + \eta \cdot \delta_j^n \cdot A_i^{n-1}.$$

输入信号从输入层经隐单元, 传向输出层, 在输出端产生输出信号, 这是工作信号的正向传播. 在信号的向前传递过程中网络的权值是固定不变的, 每一层神经元的状态只影响下一层神经元的状态. 如果在输出层不能得到期望的输出, 则转入误差信号反向传播<sup>[8]</sup>.

网络的实际输出与期望输出之间差值即为误差信号, 误差信号由输出端开始逐层向前传播, 这是误差信号的反向传播. 在误差信号反向传播的过程中, 网络的权值由误差反馈进行调节. 通过权值不断修正使网络的实际输出更接近期望输出<sup>[9]</sup>.

### 3 神经网络水源判别实验及结果分析

#### 3.1 样本设计

我们利用淮南矿业集团的谢一煤矿的水质资料进行实验. 判别因素使用了常用的地下水化学特征离子 ( $\text{Ca}^{2+}$ 、 $\text{Mg}^{2+}$ 、 $\text{K}^{+}+\text{Na}^{+}$ 、 $\text{HCO}_3^{-}$ 、 $\text{Cl}^{-}$ 、 $\text{SO}_4^{2-}$ ), 数据来源比较便捷. 一共有来自谢一煤田的 44 个水样点水质资料, 采样时间为近 20 年. 随机选 33 个作为训练样本, 11 个作为判别样本.

谢一煤矿是淮南矿业集团的一个老矿, 淮河南岸, 八公山东麓, 南与李一矿毗邻, 北与新庄孜矿相接, 最大走向长度 4400m, 面积 9.25km<sup>2</sup>(见图 3); 其主要开采煤层为 C13、B11b、B9b、B8、B7、B4b、A3、A1 煤. 目前开采到深部; 其主要充水含水层主要有煤系砂岩裂隙含水层 (煤系砂岩)、石炭系太原群灰岩溶隙含水层 (太灰)、奥陶系灰岩溶隙含水层 (奥灰); 随开采深度的增大, 煤矿开采受水害威胁越发严重; 由于煤矿在开采过程中积累了丰富的水化学资料, 因此, 开展煤矿基于水化学特征的矿井突水水源快速判别研究意义重大.



图 3 研究区及水样点位置示意图

表 1 判别样本

编号	$\text{Ca}^{2+}$ (mg/L)	$\text{Mg}^{2+}$ (mg/L)	$\text{K}^{+}+\text{Na}^{+}$ (mg/L)	$\text{HCO}_3^{-}$ (mg/L)	$\text{Cl}^{-}$ (mg/L)	$\text{SO}_4^{2-}$ (mg/L)	TDS (mg/L)	水源
34	7.41	3.16	346.69	674.24	53.18	121.83	869.39	煤系砂岩
35	3.81	2.31	294.04	638.24	46.8	59.27	725.35	煤系砂岩
36	3.81	2.31	294.04	638.24	46.08	59.27	724.63	煤系砂岩
37	1.40	1.46	233.58	444.20	42.90	52.07	553.51	煤系砂岩
38	5.21	1.34	252.43	508.88	51.05	67.91	632.38	煤系砂岩
39	63.93	30.39	48.53	389.9	28.01	30.05	395.86	太灰
40	88.78	26.62	12.42	357.56	26.94	25.93	359.47	太灰
41	85.77	28.57	17.48	363.66	28.72	30.05	372.42	太灰
42	82.96	28.93	25.99	366.1	28.22	40.34	389.49	太灰
43	101.2	31.62	17.01	381.97	48.22	36.84	425.875	奥灰
44	84.97	12.04	11.96	251.31	37.74	24.69	297.055	奥灰

#### 3.2 模型设计

对于 Elman 网络, 由样本数据和样本设计可知, 网络输入层有 33 个神经元, 输出层有 3 个神经元 (煤系砂岩 (1 0 0), 太灰 (0 1 0), 奥灰 (0 0 1)), 通过反复实验, 中间神经元为 10 个时, 网络性能最好. 训练函数为 Trainlm.

对于 BP 网络, 采用单隐层 BP 网络, 输入层为 7 个神经元, 中间层为 15 个神经元, 输出层为 3 个神经元<sup>[10]</sup>. 神经元的传递函数为 S 型正切函数和 S 型对数函数, 训练函数为 Trainlm. 取不同的中间层神经元对网络进行训练, 得到 15 个神经元是网络性能最佳, 所以 BP 网络中间层神经元的最佳数目是 15 个.

#### 3.3 训练结果

将 33 个样本分别放入两个神经网络里进行训练, 得到结果.

Elman 网络运行结果 (见图 4):

TRAINLM, Epoch 0/1000, MSE 0.33472/1e-005, Gradient 15.9764/1e-010

TRAINLM, Epoch 16/1000, MSE 1.65279e-006/1e-005, Gradient 0.00181819/1e-010

TRAINLM, Performance goal met.

BP 网络运行结果 (见图 5):

TRAINLM, Epoch 0/1000, MSE 0.405432/1e-005, Gradient 5.49302/1e-010  
 TRAINLM, Epoch 80/1000, MSE 8.6647e-006/1e-005, Gradient 0.00101882/1e-010  
 TRAINLM, Performance goal met.

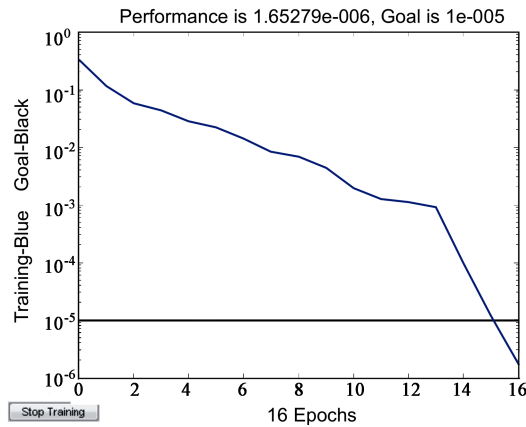


图 4 Elman 网络训练结果

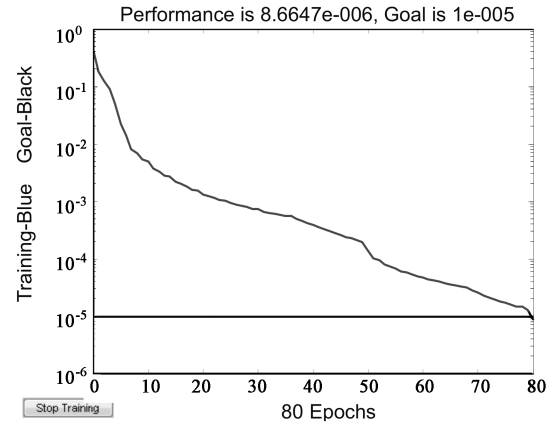


图 5 BP 网络训练结果

从图 4 可以看出, Elman 网络经过 16 次训练后, 网络误差达到要求. 从图 5 上看到, BP 网络经过 80 次训练达到了要求. 显然型前馈 BP 网络相对于反馈型 Elman 网络来说, 收敛速度要慢. 在训练过程中, BP 网络多次出现果“TRAINLM, Minimum gradient reached”, 表明 BP 网络达到梯度最小, 显然是训练陷入了局部极小点. 所以, 采用 Elman 网络进行预测, 结果更准确, 更安全.

### 3.4 测试结果

用后 11 个数据对 Elman 网络和 BP 网络进行测试, 看网络能否精确的判别水样的水源. 网络仿真函数为:  $y = \text{sim}(\text{net}, p_{\text{test}})$ . Elman 和 BP 算法判别结果见表 2.

表 2 Elman 和 BP 算法判别结果

编号	34	35	36	37	38	39	40	41	42	43	44
Elman	0.9998	0.9999	0.9999	1.0000	1.0000	0.0003	0.0001	0.0001	0.0001	0.0000	0.0001
	0.0000	0.0000	0.0000	0.0004	0.0002	0.9969	1.0000	1.0000	1.0000	0.9983	0.0000
	0.0002	0.0000	0.0000	0.0000	0.0000	0.0022	0.0000	0.0000	0.0000	0.0014	0.9999
BP	0.9997	0.9999	0.9999	0.9993	0.9997	0.0003	0.0003	0.0003	0.0002	0.0005	0.0002
	0.0068	0.0028	0.0029	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.9784	0.0000
	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0084	1.0000

从结果看, Elman 网络和 BP 神经网络都判错了 43 号水样, 将奥灰判别成了太灰. 其他十个点的判别均十分正确. 其原因是因为谢一煤煤矿的太灰和奥灰含水层的水力联系十分密切, 水质十分相似. 所取水样的地点是奥灰, 但是此处的水源可能混合了来自太灰的水.

对 Elman 网络和 BP 网络判别的误差曲线进行比较, 可以发现 Elman 比 BP 更为精确一些 (见图 6).

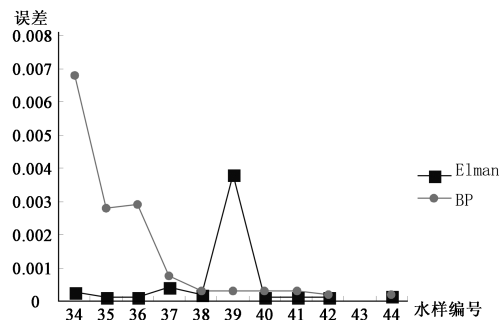


图 6 误差曲线

## 4 结论与讨论

以谢一煤矿为例, 通过将 Elman 网络与 BP 网络对比, 分别建立突水判别模型, 实例结果表明, Elman 网络模型比 BP 网络模型具有更高的判别精度, 更快的运算速度, 更好的反应地下水系统特性。

煤矿经过多年开采, 由于本身地下水的流动加上人为的干扰, 含水层水的水质在不断变化, 地下水水质是一个动态的系统。利用静态前馈 BP 网络对动态系统进行辨识时, 由于是将动态的模型问题变成静态空间建模问题, 造成网络输入结点过多、训练困难及对外部噪声敏感等弊病, 甚至更容易陷入局部极小点。相比之下, 动态回归 Elman 网络提供了一种极具潜力的选择, 它是在 BP 网络基本结构的基础上, 通过内部状态使其具备映射动态特征的功能, 从而使系统具有适应时变的能力, 因此它能够更生动、更直接地反应系统动态特性。

## 参考文献

- [1] 徐忠杰, 杨永国, 汤琳. 神经网络在矿井水源判别中的应用 [J]. 煤矿安全, 2007, 38(2): 4-6.  
Xu Z J, Yang Y G, Tang L. Application of BP neural network in evaluation of water source in mine[J]. Safety in Coal Mines, 2007, 38(2): 4-6.
- [2] 魏永强, 梁化强, 任印国, 等. 神经网络在判别煤矿突水水源中的应用 [J]. 江苏地质, 2004, 28(1): 36-38.  
Wei Y Q, Liang H Q, Ren Y G, et al. Application of neural network in discriminating water bursting source of the coal mine[J]. Jiangsu Geology, 2004, 28(1): 36-38.
- [3] 刘永霞, 冯仲科, 杜鹏志. Elman 动态递归神经网络在树木生长预测中的应用 [J]. 北京林业大学学报, 2007, 29(6): 99-103.  
Liu Y X, Feng Z K, Du P Z. Application of Elman dynamic recurrent neural network to forecast tree growth[J]. Journal of Beijing Forestry University, 2007, 29(6): 99-103.
- [4] Cheng Y L, Huang J C, Yang W C. Modeling word perception using the Elman network[J]. Neurocomputing, 2008, 71(16-18): 3150-3157.
- [5] 胡晓龙, 郜振华, 马光红. 基于 Elman 神经网络的房地产价格预测 [J]. 统计与决策, 2008, (7): 64-65.  
Hu X L, Gao Z H, Ma G H. Elman neural network based on the real estate price forecast[J]. Statistics and Decision, 2008, (7): 64-65.
- [6] Kanti K M, Rao P S. Prediction of bead geometry in pulsed GMA welding using back propagation neural network[J]. Journal of Materials Processing Technology, 2008, 200: 300-305.
- [7] Lee T L. Back-propagation neural network for the prediction of the short-term storm surge in Taichung harbor, Taiwan[J]. Engineering Applications of Artificial Intelligence, 2008, 21: 63-72.
- [8] 侯木舟, 韩旭里. 基于 MATLAB 的神经网络在股市预测中的应用 [J]. 系统工程, 2003, 21(2): 112-115.  
Hou M Z, Han X L. Stock price prediction with neural network based on MATLAB[J]. Systems Engineering, 2003, 21(2): 112-115.
- [9] 高隽. 人工神经网络原理及仿真实例 [M]. 北京: 机械工业出版社, 2003: 48-57.  
Gao J. Artificial Neural Network and the Simulation[M]. Beijing: China Machine Press, 2003: 48-57.
- [10] 李晓峰, 徐玖平, 王荫清, 等. BP 人工神经网络自适应学习算法的建立及其应用 [J]. 系统工程理论与实践, 2004, 24(5): 1-8.  
Li X F, Xu J P, Wang Y Q, et al. The establishment of self-adapting algorithm of BP neural network and its application[J]. Systems Engineering — Theory & Practice, 2004, 24(5): 1-8.