

# 优化SVM在锅炉负荷预测中的应用

陈其松<sup>1</sup>, 陈孝威<sup>1</sup>, 张欣<sup>2</sup>, 吴茂念<sup>1</sup>

(1. 贵州大学计算机科学与技术学院 贵阳 550025; 2. 贵州省光电子技术与应用重点实验室 贵阳 550025)

**【摘要】** 提出智能优化支持向量机算法来提高模型的预测能力和泛化能力。该算法针对支持向量机噪声敏感问题采用小波方法对数据集去噪; 利用核主成分分析方法提取数据特征; 采用量子粒子群算法优化支持向量机超参数。将该优化算法应用于锅炉负荷短期预测, 实验结果表明, 该优化算法预测精度较高, 收敛速度较快, 泛化性能优于其他预测方法, 且工程实现容易。

**关键词** 预测; 核主成分分析; 优化; 量子粒子群算法; 支持向量机

中图分类号 TP301.1; TP18

文献标识码 A

doi:10.3969/j.issn.1001-0548.2010.02.035

## Optimal Support Vector Machine Model for Boiler Load Forecasting

CHEN Qi-song<sup>1</sup>, CHEN Xiao-wei<sup>1</sup>, ZHANG Xin<sup>2</sup>, and WU Mao-nian<sup>1</sup>

(1. School of Computer Science and Technology, Guizhou University Guiyang 550025;

2. Guizhou Key Laboratory for Photoelectric Technology and Application Guiyang 550025)

**Abstract** Intelligently optimal support vector machine (SVM) were introduced in electric utility boiler to improve short-term load forecasting accuracy and generalization ability. Wavelet transform is adopted to filter noise in training and testing data set. Kernel principle component analysis is used in feature selection. Then quantum-behaved particle swarm algorithm is chosen to determinate optimal hyper-parameter in SVM. This optimal algorithm has been tested on power plant and the results show that the prediction can get higher precision and convergence speed.

**Key words** forecasting; kernel principle component analysis; optimization; quantum-behaved particle swarm algorithm; support vector machines

建立在统计学习理论基础上的支持向量机<sup>[1-3]</sup>采用结构风险最小化原则设计学习机器<sup>[4-5]</sup>, 综合考虑经验风险和模型的复杂性, 较好地解决了过学习、非线性、高维数等问题。由于其最终可以归结为求解一个凸二次规划问题, 获得的解是全局最优解, 避免了神经网络等方法的局部极值问题, 因此具有较好的推广能力。随着应用的广泛, 支持向量机的不足也逐渐被关注, 主要体现在: (1) 由于标准支持向量机的数学模型是一个二次规划问题, 使得对于大规模样本问题的训练速度较慢。(2) 在小规模样本条件下, 模型参数取值缺少先验信息, 搜索区间较大, 在多个参数需要优化的情况下, 优化过程花费较多的时间。(3) 在模型参数选择方面, 目前还缺乏有效的方法和理论依据。(4) 多类分类问题。(5) 噪声敏感问题。针对支持向量机的不足, 出现了多种改进方法<sup>[6]</sup>, 如解决支持向量机训练速度问题的分块分解算法、序贯最小优化算法、模型参数选择的

遗传算法、蚁群算法等。但在实际应用中, 仅对支持向量机算法的某个方面进行改进是不够的。本文针对支持向量机噪声敏感问题采用小波方法先对样本去噪, 再采用核主成分分析方法提取样本数据特征, 超参数的寻优采用粒子群算法完成, 构建了智能优化支持向量机方法。

## 1 智能优化支持向量机方法

### 1.1 支持向量机<sup>[1-3]</sup>

给定 $l$ 个训练样本 $\{x_i, y_i\}_{i=1}^l$ , 集合 $\{x_i\} \in R^n$ 为 $n$ 维训练样本输入,  $y_i \in \{+1, -1\}$ 为训练样本输出。若样本线性可分则可由超平面 $\omega X + b = 0$ 分开, 使得每个样本点都满足:

$$y_i(x_i \omega + b) - 1 \geq 0 \quad i = 1, 2, \dots, l \quad (1)$$

此时分类间隔为 $2/\|\omega\|$ 。所以满足式(1), 且使 $\|\omega\|^2$ 最小的分类面就是最优分类面。

利用Lagrange优化方法, 最优分类面问题转换

收稿日期: 2008-10-06; 修回日期: 2009-05-12

基金项目: 贵州省自然科学基金(黔科合字20072004); 贵州省省长专项资金(黔省专合字(2007)14号)

作者简介: 陈其松(1974-), 男, 博士生, 副教授, 主要从事信号处理、自动控制、网络与通讯技术等方面的研究。

为求解如下的QP问题:

$$\min W(\alpha) = 1/2 \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^l \alpha_i \quad (2)$$

$$\text{s.t. } \sum_{i=1}^l y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i=1,2,\dots,l \quad (3)$$

式中  $\alpha_i$  是QP优化问题所求解的Lagrange乘子;  $C$  为惩罚参数;  $k(x_i, x_j)$  为核函数。常用的核函数有多项式核、高斯径向基核、Sigmoid核等。

假设最优解是  $\alpha^* = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ , 则最终所得到的决策函数为:

$$f(x_i) = \sum_{i=1}^l y_i \alpha_i^* k(x, x_i) + b \quad (4)$$

式中  $f(x_i)$  为SVM相对于第  $i$  个样本的输出。 $f(x_i)$  值的正负表明归属类型, 故训练一个支持向量机就是解决一个线性约束的二次规划问题。

### 1.2 小波阈值去噪

基于小波变换<sup>[7]</sup>的多分辨分析技术在信号去噪技术中有明显的优点。对含噪信号进行小波变换后得到对应信号的小波系数幅度较大, 数目较少, 而对应噪声的小波系数幅值较小。通过在不同尺度上选取合适的阈值, 并将小于该阈值的小波系数置零, 而保留大于阈值的小波系数, 可滤除信号中的噪声。最后进行小波逆变换, 得到的重构信号即为经过滤波的信号。

### 1.3 核主成分分析方法

高维数据集会导致支持向量机核矩阵的计算量很大, 造成维数灾难。特征提取的主成分分析(principle component analysis, PCA)方法就是常用的对高维数据降维以获得数据主要信息的方法<sup>[8]</sup>。对训练样本集  $\mathbf{X} = \{x_1, x_2, \dots, x_M\}$  求其协方差矩阵  $\mathbf{C}_x$  (假设各个行向量的均值皆为0), 则有:

$$\mathbf{C}_x = \mathbf{X}\mathbf{X}^T = \mathbf{U}\mathbf{A}\mathbf{U}^T \quad (5)$$

式中  $\mathbf{A} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_M]$ ,  $\lambda_i$  称为特征值;  $\mathbf{U}$  的每一列称为一个特征矢量。式(5)即是对  $\mathbf{C}_x$  的主分量分解, 分解出的分量是按能量大小排序的。如果原始数据的秩小于  $M$ , 则某些奇异值和特征值等于0。因此, PCA是一种数据压缩和特征提取的有效途径。

由于变量之间的关系往往是非线性的, 常规PCA不再适用, 而基于核的主成分分析(kernel principle component analysis, KPCA)方法作为PCA方法的非线性扩展, 可以有效地提取输入数据的非线性信息<sup>[9]</sup>。KPCA通过一个非线性映射  $\Phi$  将样本映

射到特征空间  $F$ , 即:

$$\Phi: x \rightarrow \Phi(x) \in F \quad (6)$$

然后对数据集  $\Phi(\mathbf{X}) = \{\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)\}$  进行PCA。假设核样本已经中心化, 即  $\sum_{i=1}^M \Phi(x_i) = 0$ , 其协方差矩阵为:

$$\mathbf{C} = \frac{1}{M} \sum_{i=1}^M \Phi(x_i) \Phi(x_i)^T \quad (7)$$

解  $\mathbf{C}$  的特征值  $\lambda \geq 0$  及相应特征向量  $\mathbf{v} \in F$ , 即:

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} \quad (8)$$

由再生核理论, 存在一组系数  $\alpha_i (i=1, 2, \dots, M)$  使:

$$\mathbf{v} = \sum_{i=1}^M \alpha_i \Phi(x_i) \quad (9)$$

定义  $M \times M$  的核矩阵  $\mathbf{K} = \langle \Phi(x_i) \cdot \Phi(x_j) \rangle$ , 结合式(8)和式(9)可得:

$$\mathbf{K}\alpha_i = M\lambda\alpha_i \quad i=1, 2, \dots, M \quad (10)$$

### 1.4 量子行为粒子群算法

优化粒子群(particle swarm optimization, PSO)算法的基本思想<sup>[10]</sup>是: 将优化问题的每一个解称为粒子, 每个粒子根据自己和其他粒子的“飞行经验”群游, 从而达到从全空间搜索最优解的目的。每个粒子表示在  $n$  维空间中的一个点, 用  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]$  表示第  $i$  个粒子, 第  $i$  个粒子的个体最优解  $\text{pbest}_i = [p_{i1}, p_{i2}, \dots, p_{in}]$ ; 全局最优解  $\text{gbest}_i = [g_1, g_2, \dots, g_n]$ , 而  $x_i$  的第  $k$  次迭代修正量(粒子移动的速度)  $v_i^k = [v_{i1}^k, v_{i2}^k, \dots, v_{in}^k]$ 。每个粒子根据式(11)、式(12)更新自己的速度和位置:

$$v_{id}^k = w_i v_{id}^{k-1} + c_1 \text{rand}_1(p_{id}^{k-1} - x_{id}^{k-1}) + c_2 \text{rand}_2(g_d^{k-1} - x_{id}^{k-1}) \quad (11)$$

$$x_{id}^k = x_{id}^{k-1} + v_{id}^k \quad (12)$$

式中  $k$  表示第  $k$  次迭代;  $i=1, 2, 3, \dots, m$ ;  $d=1, 2, 3, \dots, n$ ;  $m$  为粒子群中粒子的个数;  $n$  为解向量的维数;  $c_1$  和  $c_2$  为加速度因子, 分别是两个正常数;  $\text{rand}_1$  和  $\text{rand}_2$  是两个独立的介于[0,1]之间的随机数;  $w_i$  为动量项系数, 调整其大小可改变搜索能力的强弱。

量子粒子群算法(QPSO)是文献[11]从量子力学的角度, 通过对粒子收敛行为的研究得到的一种新的PSO算法。该算法将每个个体看作是在  $n$  维搜索空间中的一个没有重量和体积的微粒, 并在搜索空间中以一定的速度飞行。每个微粒代表  $n$  维空间中的一个位置, 朝以下两个方向调整微粒的位置:

(1) 至今发现的每个微粒的最优位置。(2) 微粒

群的最优位置。

每一个微粒  $i$  包含下列信息: (1) 微粒的当前位置  $x_i = [x_{i1}, x_{i2}, \dots, x_{id}]$ 。(2) 微粒  $i$  的最佳适应性值  $pbest: P_i = [P_{i1}, P_{i2}, \dots, P_{id}]$ 。(3) 微粒群的最佳适应性值  $gbest: P_g = [P_{g1}, P_{g2}, \dots, P_{gd}]$ 。

在QPSO算法中, 每一个粒子的移动方式为:

$$mbest = (1/M) \sum_{i=1}^M P_i = [(1/M) \sum_{i=1}^M P_{i1}, P_{i2}, \dots, P_{ij}] \quad (13)$$

$$pp_{ij} = \varphi P_{ij} + (1-\varphi) P_{gj}, \quad \varphi = \text{rand}(0,1) \quad (14)$$

$$x_{ij} = pp_{ij} + \alpha |mbest_j - x_{ij}| \ln(1/u), \quad u = \text{rand}(0,1) \quad (15)$$

式中  $mbest$  是粒子群  $pbest$  的中间位置;  $pp_{ij}$  为  $P_{ij}$  和  $P_{gj}$  之间的随机点;  $\alpha$  为QPSO的收缩扩张系数, 它是QPSO收敛的一个重要参数。在QPSO算法中, 粒子的状态用位置向量来描述, 且只有一个控制参数  $\alpha$ 。对该参数的选择和控制是非常重要的, 它关系到整个算法的收敛速度<sup>[12]</sup>。

## 2 智能优化SVM算法用于锅炉负荷预测

### 2.1 系统工况介绍

电站锅炉主蒸汽温度具有大延时特性。传统的PID控制是一种固定参数控制, 难以同时使稳定性和控制品质都取得良好的效果。本文将智能优化SVM算法应用于300 MW电力锅炉的短期负荷预测, 根据预测结果调节控制器, 保持过热蒸汽的出口温度。

原锅炉主汽温自动调节系统采用串级调节, 方框图如图1所示。

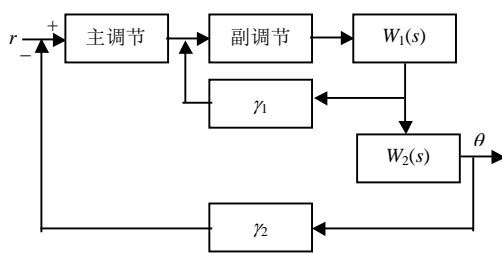


图1 主汽温串级调节系统框图

$$\text{图中: } W_1(s) = \frac{9}{1+87s} e^{-54s}, \quad W_2(s) = \frac{2}{1+37s} e^{-7s},$$

$$\gamma_1 = \gamma_2 = 0.1 \text{ mA}/^\circ\text{C}。$$

### 2.2 数据预处理

本文采用小波变换对进行训练和预测的锅炉负荷数据进行滤波去噪。

首先对历史负荷数据中包含的缺失和错误数据预处理, 然后采用小波变换<sup>[13-14]</sup>对数据进行滤波去

噪。本文采用db4小波和sqtwolog的阈值形式进行3层分解, 阈值为 $\sqrt{2\lg(\text{length}(X))}$ 。

### 2.3 KPCA特征提取

选择KPCA的核函数和核参数, 并进行数据的特征选择, 将提取到的非线性主成分作为SVM的输入数据。通常  $\sum_{i=1}^M \Phi(x_i) = 0$  不一定成立, 所以将核矩阵换成:

$$\bar{K} = K - I_M \cdot K - K \cdot I_M + I_M \cdot K \cdot I_M \quad (16)$$

式中  $I_M$  为  $M \times M$  的矩阵, 且  $(I_M)_{i,j} = \frac{1}{M}$ 。

核函数选用与支持向量机相同的高斯径向基核函数  $k(x, y) = \exp(-\|x - y\|^2 / \sigma^2)$ , 核参数  $\sigma$  取0.1, 误差惩罚参数  $C$  取3, 训练样本数取72, 贡献比为84.74%。

### 2.4 SVM核函数选择和超参数优化<sup>[15-16]</sup>

本文中支持向量机核函数选择高斯径向基核函数, 利用QPSO算法对核参数  $\sigma$  和惩罚系数  $C$  进行优化选择, 选用  $k$ -fold交叉验证误差作为SVM参数选择的目标值。

算法具体步骤如下:

1) 读取样本数据, 随机产生  $\{C, \sigma\}$  作为粒子的初始位置。

2) 把整个样本平均分成  $k$  个互不包含的子集  $S_1, S_2, \dots, S_k$ 。

3) 根据当前的  $\{C, \sigma\}$  训练SVM, 计算  $k$ -fold交叉验证误差; (1) 初始化  $i=1$ ; (2)  $S_i$  子集留作检验集, 其余的子集合并作为训练集, 训练SVM; (3) 计算第  $i$  个子集的泛化误差  $e_i = \text{mean}(S_i - \hat{S}_i)^2$ , 令  $i=i+1$ , 重复步骤(2)直到  $i=k+1$ ; (4) 计算  $k$  次泛化误差的平均值得到  $k$ -fold交叉验证误差。

4) 以  $k$ -fold交叉验证误差作为适应值, 并记忆个体与群体所对应的最佳适应值的位置为  $pbest$  和  $gbest$ , 根据式(11)和式(12)搜寻更好的  $\{C, \sigma\}$ 。

5) 重复步骤2)直到满足最大迭代次数。

6) 结束。

选择结果是核参数  $\sigma$  取3.99, 误差惩罚参数  $C$  取237。

### 2.5 SVM的训练

求解二次规划问题是SVM训练算法最核心、最主要的问题。将QP问题分解为若干个规模一定的子QP问题, 用QPSO优化SVM子问题。分解QP问题即选择一个工作集作为优化的对象。从  $L$  个训练样本

点中选择  $q(q < l)$  个样本点所对应的Lagrange乘子  $\alpha_i$  作为工作集  $\alpha_w$ , 其余样本点所对应的Lagrange乘子  $\alpha_i$  保持不变, 作为非工作集  $\alpha_N$ 。用分解算法选择出工作集和非工作集后, 对工作集在可行域里用QPSO算法对QP问题求解最优值。粒子数一般选择在20~40个粒子之间, 使所有的粒子初始化时都满足  $\alpha_w^T y_w + \alpha_N^T = 0$ , 即粒子的飞行方向满足可行方向法。算法如下:

- (1) 初始化粒子群, 对每个粒子初始化一个满足约束条件的  $\alpha_w$  值。
- (2) 计算粒子的目标函数值。
- (3) 更新每个粒子  $\alpha_w$  的局部最优值。
- (4) 更新  $\alpha_w$  全局最优值。
- (5) 根据式(13)计算 mbest。
- (6) 根据式(14)计算每个粒子随机点  $pp_{ij}$ 。
- (7) 根据式(15)更新每个粒子的  $\alpha_w$  值, 使拉格朗日乘子的值在  $[0, C]$  内,  $C$  为惩罚参数。
- (8) 返回第(2)步, 直到满足迭代的次数。
- (9) 结束。

### 3 实验结果分析

本文将样本分为288组训练验证集和72组测试集, 其中测试集取连续三天的数据。将训练验证集随机均分为4组, 用3组数据训练, 1组数据验证。程序运行计算至迭代结束, 搜索得到支持向量机的选定惩罚因子  $C$  为237, 高斯核函数参数  $\sigma$  为3.99。预测结果如图2所示。

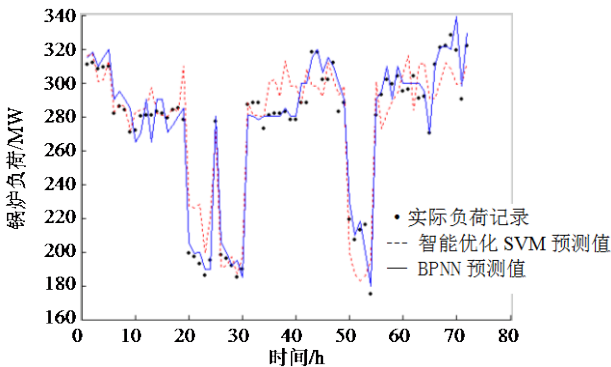


图2 锅炉负荷实测值与预测值的比较

为体现本文的智能优化SVM模型的预测性能, 将其与BP网络对比研究。BP网络采用相同的数据集, 在BP网络训练前对训练集数据和测试集数据采用归一化, 并在模型预测后对预测结果采用反归一化, 数据处理的其他环节与本文采用的智能优化SVM模型相同。

为便于比较, 分别利用遗传算法(交叉概率为0.8, 变异概率为0.2%, 种群数和最大迭代次数与QPSO法一样)和正交法(2因素10水平)进行多次试验, 对所得SVM模型进行测试, 平均结果如表1所示。

从表中可见, 本文采用的方法测试均方误差(MSE)最小, 且支持向量个数(SV)最少。

表1 SVM参数选择试验结果

优化方法	$C$	$\sigma$	SV	训练误差	测试误差
GA5-fold	261	3.71	32	0.007 8	$8.43 \times 10^{-4}$
正交法5-fold	164	2	28	0.008 1	$1.24 \times 10^{-3}$
QPSO5-fold	237	3.99	26	0.007 4	$6.12 \times 10^{-4}$

对比本文智能优化SVM预测模型和BP网络预测模型的相对误差如表2所示。

表2 智能优化SVM预测模型和BP网络预测模型的相对误差对比

相对误差/(%)	误差分布/(%)	
	智能优化SVM模型	BP网络模型
0~5	86.15	56.17
5~10	12.56	28.54
10~15	1.29	15.29

### 4 总 结

本文提出的智能优化SVM算法采用小波去噪方法解决SVM的噪声敏感问题; 采用KPCA提取数据特征的方法解决样本维数灾难问题; 利用QPSO对SVM的超参数寻优克服了试算的低效率和盲目性。这些措施显著提高了SVM的学习性能。将该优化算法应用于火电锅炉的短期负荷预测系统, 取得了良好的效果。

### 参 考 文 献

- [1] VAPNIK V N. An overview of statistical learning theory[J]. IEEE Trans Neural Network, 1999, 10(5): 988-999.
- [2] SHI Guang-zhi, HU Jun-chuan. A dynamic recognition method study using the support vector machine[C]//Mechatronics and Automation, ICMA 2007 International Conference on Piscataway. NJ: IEEE Press, 2007: 1694-1698.
- [3] SANCHEZ A D. Advanced support vector machines and kernel methods[J]. Neurocomputing, 2003, 55(1): 5-20.
- [4] ZHANG Jun-feng, HU Shou-song. Nonlinear time series fault prediction based on clustering and support vector machines[J]. Control Theory & Applications, 2007, 24(1): 64-68.
- [5] SUN Da-ru, WU Le-nan. Face recognition based on nonlinear feature extraction and SVM[J]. Journal of Electronics & Information Technology, 2004, 26(2): 307-311.
- [6] JOHN C P. Fast training of support vector machines using sequential minimal optimization[C]//Advances in Kernel Methods-Support Vector Learning. MA: MIT Press, 1999: 185-208.
- [7] 程正兴. 小波与小波变换导论[M]. 北京: 机械工业出版社

- 社, 2008.
- CHEN Zheng-xin. An introduction to wavelets[M]. Beijing: China Machine Press, 2008.
- [8] CAO L J, CHUA K S, CHONG W K. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine[J]. Neurocomputing, 2003, 55(2): 321-336.
- [9] SCHOLKOPF B, SMOLA A J, MULLER K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation, 1998, 10(6): 1299-1319.
- [10] 沈 艳, 郭 兵, 古天祥. 粒子群优化算法及其与遗传算法的比较[J]. 电子科技大学学报, 2005, 34(5): 696-699.
- SHEN Yan, GUO Bing, GU Tian-xiang. Particle swarm optimization algorithm and comparison with genetic algorithm[J]. Journal of University of Electronic Science and Technology of China, 2005, 34(5): 696-699.
- [11] SUN J, FENG B, XU W B, et al. Particle swarm optimization with particles having quantum behavior[C]// IEEE Proc of Congress on Evolutionary Computation. Portland, United States: IEEE, 2004: 619-623.
- [12] VLACHOGIANNIS J G, LEE KY. A comparative study on particle swarm optimization for optimal steady-state performance of power systems[J]. Power Systems, IEEE Transactions on, 2006, 4(21): 1718-1728.
- [13] HO K C, PROKOPIW W, CHAN Y T. Modulation identification of digital signals by the Wavelet transform[J]. IEEE Proc-Radar, Sonar Navig, 2000, 147(4): 169-176.
- [14] SATISH L, NAZNEEN B. Wavelet-based denoising of partial discharge signals buried in excessive noise and interference[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2003, 10(2): 354-367.
- [15] CHERKASSKY V, MA Y. Practical selection of SVM parameters and noise estimation of SVM regression[J]. Neural Networks, 2004, 17(1): 113-126.
- [16] CHAPPELLE O, VAPNIK V, BOUSQUET O, et al. Choosing multiple parameters for support vector machines[J]. Machine Learning, 2002, 46(1): 131-160.

编辑 漆蓉

· 我校科研成果介绍 ·

### 异种机网络消息收发系统Mesman

Mesman是一个基于异种机网络环境的消息收发系统,它采用客户/服务器体系结构,消息服务器负责消息(文件)的存储、转发和日志管理,消息客户则提供消息的发送、接收和实时提示功能。该系统为PC机,工作站X终端的用户提供基于文件传送的TCP/IP网络消息通信机制,使网络上各席位间能够迅速、准确和方便地传送各类信息,并确保对所有信息的集中管理。Mesman的总体技术指标如下:

- (1) 运行环境: Sun工作、X-Tenminal、386或以上PC机;
- (2) 操作系统: Unix、Windows;
- (3) 网络规模: 大中小型局域网、广域网;
- (4) 传送协议: TCP/IP;
- (5) 消息系数: 仅受服务器硬盘空间的限制。