基于相关向量机的小世界神经元网络拓扑估计*

郝崇清，王 江，邓 斌，魏熙乐
（天津大学 电气与自动化工程学院，天津 300072）

摘 要：采用相关向量机从含噪时间序列中估计小世界神经元网络的节点动力学方程和拓扑结构。在具有多项式结构或能以幂级数展开的动力学系统中，将未知方程写成统一多项式形式，原动力学方程的项在统一多项式中是稀疏的，利用稀疏贝叶斯学习估计出稀疏项从而实现动力学方程和拓扑结构的估计。利用该方法对 FHN 小世界神经元网络进行节点动力学方程和拓扑估计，结果表明，该方法能快速准确地估计节点动力学方程结构和网络拓扑，对动力学方程系数和网络耦合强度有很高的估计精度，而且对噪声有强鲁棒性。

关键词：小世界网络；相关向量机；动力学方程重建；拓扑估计；神经元模型
中图分类号：TP274；TP393 **文献标志码**：A **文章编号**：1001-3695(2012)11-4082-03
doi:10.3969/j.issn.1001-3695.2012.11.020

Estimation of small-world neuronal network topology based on relevance vector machine

HAO Chong-qing, WANG Jiang, DENG Bin, WEI Xi-le
(*School of Electrical & Automation Engineering, Tianjin University, Tianjin* 300072, *China*)

Abstract: This paper applied relevance vector machine to estimate node dynamical equations and topology of small-world neuronal network from noisy time series. According to the fact that many dynamical equations or its power series expansion had polynomial structure, by constructing a unified polynomial and making the original dynamical equation sparse in the unified polynomial, obtained dynamical equations and network topology obtained while used sparse Bayesian learning to estimate the sparse nonzero terms. FHN small-world neuronal network as a paradigm demonstrated the estimation effect of the dynamical equations and network topology. The results show that the estimating strategy can identify equations structure and network topology accurately and quickly, the error is small in dynamical equations coefficients and couple strength estimation and is robust to noise.

Key words: small-world networks; relevance vector machine (RVM); dynamical equations reconstruction; topology estimation; neuron model

0 引言

对于一个非线性动力学系统，当给定系统结构的情况下估计系统的动力学参数，常用的方法有自适应同步、状态估计等[1,2]。由非线性动力学系统构成的复杂网络中，网络节点的动力学和网络拓扑决定了网络的宏观行为，研究网络节点的不同动力学和不同拓扑情况下网络的动力学演化机制是近来研究的热点[3]。然而网络节点的动力学和网络拓扑往往是未知的，动力学未知包括方程结构和参数都未知，这为深入研究网络的动力学行为带来了困难，因此，估计和辨识网络节点方程和拓扑结构具有重要意义。

由系统的输出时间序列估计和辨识节点动力学和拓扑是动力学研究的逆问题。对于多项式结构或者可以通过幂级数展开成多项式结构的动力学系统，首先指定系统方程的统一形式，即将方程的全局重建变成一个线性回归问题，利用回归系数与结构项的对应关系从而同时得到方程的结构和参数。Gouesbet 等人[4]将未知方程写成自治多项式结构的统一形式进行估计；Lu 等人[5]给出了含有观测噪声的自治多项式结构的估计；Bezruchko 等人[6]提出非自治微分方程结构的估计方法，其多是利用最小二乘法估计多项式系数，然后利用多项式系数与方程结构的对应关系得到系统的方程。然而这种估计方法收敛速度慢且统一多项式的零项与非零项均存在较大稳态误差。为了预测系统的突变及网络动力学特性，Wang 等人[7,8]提出了一种基于压缩传感的动力系统方程估计和网络拓扑结构估计方法，但该方法对噪声的鲁棒性不强。

由于统一结构的多项式项中非零项个数远大于零项，所以这是一种稀疏结构的多项式形式；而相关向量机是求解噪声环境下此种结构稀疏解的有效方法，即映射到以稀疏贝叶斯学习为基础的相关向量机（RVM）[9]框架下进行估计。相比于支持向量机（SVM）结构风险最小化的学习机，RVM 是基于贝叶斯框架下的学习机，由于不受 Mercer 条件的限制，可以构建任意核函数，通过假设权值服从均值为零的正态分布，将权值过小

收稿日期：2012-04-28；**修回日期**：2012-05-31 **基金项目**：国家自然科学基金资助项目(61072012,61172009)；国家自然科学基金青年基金资助项目(60901035)

作者简介：郝崇清(1981-)，男，山东枣庄人，博士研究生，主要研究方向为复杂网络、生物电信息处理等(haochongqing@tju.edu.cn)；王江(1964-)，男，教授，博导，博士，主要研究方向为神经系统非线性动力学、神经电信息处理；邓斌(1979-)，男，副教授，博士，主要研究方向为非线性系统分析与控制、生物电工学；魏熙乐(1975-)，男，副教授，博士，主要研究方向为生物电效应、神经系统建模.

的项置为零,克服了 SVM 容易过拟合的缺点。

1 网络动力学方程和拓扑重建

假定网络节点动力学方程形式如下:

$$\dot{x} = F(x,A)$$

其中:$x=(x_1,x_2,\cdots,x_n)\in \mathrm{R}^n$ 为系统状态变量;A 为系统参数;$F(x,A)=(F_1(x,A_1),F_2(x,A_2),\cdots,F_n(x,A_n))$ 为平滑向量函数。假设系统各变量均可观测,则每个系统变量可以表示成统一形式:

$$\dot{x}_i = \sum_{l_1,l_2,\cdots,l_d=0}^{N}(a_i)_{l_1,l_2,\cdots,l_n}\prod_{j=1}^{d}x_j^{l_j}$$

其中:d 为变量的最高幂指数,系数值 $(a_i)_{l_1,l_2,\cdots,l_d}$ 为向量 $\boldsymbol{A}_i$ 的元素,写成向量形式为 $\dot{x}_i=F_i(x,\boldsymbol{A}_i)=\varphi A_i^{\mathrm{T}}$,$\varphi=\sum_{l_1,l_2,\cdots,l_d=0}^{N}\prod_{j=1}^{d}x_j^{l_j}$。选择适当的幂指数 d 时,由于缺少相关项,$\boldsymbol{A}_i$ 中的大部分元素为零,即其是稀疏的。当系统在某参数 A 下演化时,获得 n 个系统状态变量的 μ 个测量值 $[x(t_1),x(t_2),\cdots,x(t_\mu)]^{\mathrm{T}}$,对于第 i 个方程 F_i 由观测量 $[x_i(t_1),x_i(t_2),\cdots,x_i(t_\mu)]^{\mathrm{T}}$ 和采样频率求得

$$\dot{\chi}_i=[\dot{x}_i(t_1),\dot{x}_i(t_2),\cdots,\dot{x}_i(t_\mu)]^{\mathrm{T}}$$

同时求得矩阵 Φ 的 μ 个行向量 $\Phi=[\varphi(t_1),\varphi(t_2),\cdots,\varphi(t_\mu)]^{\mathrm{T}}$,由此获得如下等式:

$$\dot{\chi}_i=\Phi A_i^{\mathrm{T}}+e \tag{1}$$

其中:$\boldsymbol{A}_i$ 为由系数 $(a_i)_{l_1,l_2,\cdots,l_d}$ 组成的向量;e 为测量噪声。利用相关向量机可以估计出 $\boldsymbol{A}_i$ 的值。由于 $\boldsymbol{A}_i$ 与系统各项存在对应关系,估计出系统每个方程 $\boldsymbol{A}_i$ 的值便可估计出整个系统的方程。

当由 K 个节点组成网络时,为简单起见,假设节点间仅通过第一个变量存在耦合,对第 i 个节点 $[\dot{x}_1]_i$ 表示成

$$[\dot{x}_1]_i=[F_1(x,A_1)]_i+\sum_{j=1,j\neq i}^{K}C_{i,j}([x_1]_j-[x_1]_i) \tag{2}$$

其中:$i=1,2,\cdots,K$,$C_{i,j}$ 为第 i 个节点与第 j 个节点的耦合强度。对式(2)所包含的节点变量进行分组,把含有第 i 个节点的所有变量分成一组,这样可将式(2)分成 K 组,即写成如下 K 组式子的和:

$$\begin{cases}[F_1(x,A)]_i-\sum_{j=1,j\neq i}^{K}C_{i,j}[x_1]_i\\ C_{i,1}[x_1]_1\\ C_{i,K}[x_1]_K\end{cases} \tag{3}$$

对其中的每一组均可按照前面介绍的方法进行估计方程结构和参数,利用第一组和其他变量的估计可得出第 i 个节点的方程,其余各组估计出与第 i 个节点相连的节点号及耦合强度,对网络中所有节点均执行该估计方法,便得到网络的拓扑结构。

2 相关向量机方法

将式(1)重新写为如下形式:

$$y=\Phi x+e \tag{4}$$

其中:$e\sim N(0,\sigma^2 I_N)$;RVM 根据 x 在某基下是稀疏的、数据 y 是可测量的等先验知识,估计 x 的后验密度函数。K 次测量下的似然函数可以写为

$$P(y|x,\sigma^2)=(2\pi\sigma^2)^{-K/2}\exp(-\|y-\Phi x\|^2/2\sigma^2) \tag{5}$$

式(5)可以通过最大似然估计求解 x,但该方法容易引起过学习,本文引入基于稀疏贝叶斯学习的 RVM 方法求解该问题,该方法首先对稀疏向量 x 的每一个元素赋予一个零均值的高斯先验模型

$$P(x|\alpha)=\prod_{i=1}^{N}N(x_i|0,\alpha_i^{-1}) \tag{6}$$

其中:α_i 为高斯密度函数的逆方差;α 称为超参数;σ^2 也可称为超参数。对 α 和噪声方差 $\sigma^2=\beta^{-1}$ 定义一个 Gamma 分布的超先验,即

$$P(\alpha|a,b)=\prod_{i=1}^{N}\mathrm{Gamma}(\alpha_i|a,b)$$

$$P(\beta|c,d)=\prod_{i=1}^{N}\mathrm{Gamma}(\beta_i|c,d)$$

由此得到 x 的最终先验为

$$P(x|a,b)=\prod_i^N\int_0^\infty N(x_i|0,\alpha_i^{-1})\mathrm{Gamma}(\alpha_i|a,b)\mathrm{d}\alpha_i$$

积分 $\prod_i^N\int_0^\infty N(x_i|0,\alpha_i^{-1})\mathrm{Gamma}(\alpha_i|a,b)\mathrm{d}\alpha_i$ 为 t 分布,通过选择恰当的 a 和 b,t 分布在 $x_i=0$ 附近产生尖峰,因此先验使得由上式描述的先验是稀疏先验。在已知 y、α 和 σ^2 的情况下,

$$P(x|y,\alpha,\sigma^2)=\frac{P(x,y,\alpha,\sigma^2)}{P(y,\alpha,\sigma^2)}=\frac{P(y|x,\alpha,\sigma^2)P(x,\alpha,\sigma^2)}{P(y|\alpha,\sigma^2)P(\alpha,\sigma^2)}=$$

$$\frac{P(y|x,\alpha,\sigma^2)P(x|\alpha,\sigma^2)P(\alpha,\sigma^2)}{P(y|\alpha,\sigma^2)P(\alpha,\sigma^2)}=\frac{P(y|x,\sigma^2)P(x|\alpha)}{P(y|\alpha,\sigma^2)}$$

$$P(y|x,\sigma^2)P(x|\alpha)/\int P(y|x,\sigma^2)P(x|\alpha)\mathrm{d}x \tag{7}$$

将式(5)和(6)代入式(7),积分后得 $P(x|y,\alpha,\sigma^2)$ 服从多元高斯分布 $N(\mu,\Sigma)$,其中:

$$\mu=\sigma^{-2}\Sigma\Phi^{\mathrm{T}}y$$

$$\Sigma=(\sigma^{-2}\Phi^{\mathrm{T}}\Phi+A)^{-1}$$

$$A=\mathrm{diag}(\alpha_0,\alpha_1,\cdots,\alpha_N)$$

下面对 α 和 σ^2 进行估计,由于

$$P(\alpha,\sigma^2|y)\propto P(y|\alpha,\sigma^2)P(\alpha)P(\sigma^2)$$

求得 α 和 σ^2 的最大后验密度,即

$$(\alpha_{MP},\sigma^2_{MP})=\arg\max_{\alpha,\sigma^2}P(y|\alpha,\sigma^2)$$

而

$$P(y|\alpha,\sigma^2)=\int P(y|x,\sigma^2)P(x|\alpha)\mathrm{d}x=$$

$$(2\pi)^{-N/2}|\Omega|^{-1/2}\exp(-y^{\mathrm{T}}\Omega^{-1}y/2) \tag{8}$$

其中:$\Omega=\sigma^2 I_N+\Phi A^{-1}\Phi^{\mathrm{T}}$。

对式(8)取对数

$$\ln P(y|\alpha,\sigma^2)=-(N\ln 2\pi+\ln|\Omega|+y^{\mathrm{T}}\Omega^{-1}y)/2 \tag{9}$$

利用 EM 算法对式(9)进行迭代求解

$$\alpha_i^{\mathrm{new}}=\gamma_i/\mu_i$$

$$(\sigma^2)^{\mathrm{new}}=\|y-\Phi\mu\|^2/N-\sum_{i=0}^{N}\gamma_i$$

其中:$\gamma_i=1-\alpha_i\Sigma_{i,i}$,$\Sigma_{i,i}$ 为 Σ 中对角线上的第 i 个值。给定 α 和 σ^2 的初值,可以通过迭代更新就能逼近 α_{MP} 和 σ^2_{MP}。得到 x 的后验密度函数后,取后验估计的均值为 x 的点估计,迭代若干次后大部分 α_i 趋于无穷大,即对应的 x_i 趋于零,其他趋于有限值 x_i 就称之为相关向量。

3 小世界神经元网络拓扑估计

考虑 FHN 神经元模型作为网络节点,节点间通过膜电压形成 gap 连接构造成一个小世界网络,其数学描述为

$$\dot{x}_i=100x_i-33.33x_i^3-100y_i+\sum_j\theta_{ij}(x_j-x_i)$$

$$\dot{y}_i=x_i+0.2 \tag{10}$$

其中:θ_{ij} 为第 i 个和第 j 个节点的耦合强度,当 i、j 间有连接时 $\theta_{ij}=1$,否则为零。小世界网络的最邻近边数为 $k=4$,重连概率为 $p=0.1$,网络节点数为 20,网络类型为 NW 型小世界网络[10],即以随机加边的方法构造的网络模型。网络拓扑结构如图 1 所示。

利用输出时间序列估计网络拓扑时，统一多项式的最高幂指数取为3，式(10)的第一个等式写成统一多项式时共有320项，非零项个数很少向量显然是稀疏的，即使是各节点全连接时仍是稀疏的。当各节点的输出时间序列中含有强度为0.002的高斯噪声且采样点数为3 000时，网络中第12个节点N_{12}处统一多项式系数估计值如图2所示。图2中三个较大的值为式(10)的第一个方程的系数估计值，其中－100.0472对应$-100y_{12}$项，－33.3579对应$-33.33x_{12}^3$项。由于N_{12}与其他六个节点相连，所以式(10)的第一个方程中如果将x_{12}对应的项进行合并则应为$94x_{12}$，其估计值为94.225。以同样方式估计式(10)的第二个方程，便得到了节点的动力学方程以及与其耦合的节点。由于系数值与耦合强度数值上差异很大，所以将三个大的系数值置为零后重新显示于图2左上方。

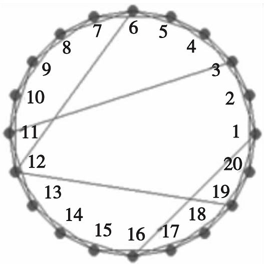

图1　FHN神经元小世界网络拓扑

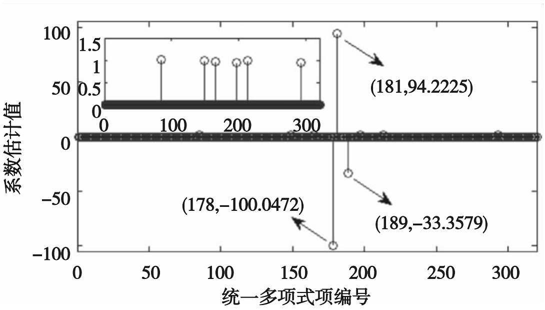

图2　节点N_{12}第一个方程统一多项式系数估计值

虽然式(10)中第二个方程并没有与其他节点耦合，如果仍将其写成第一个方程的统一多项式形式显然也是可以的，并且非零项非常稀疏，这里考虑相对不稀疏的情况，即不考虑耦合的情况，此时统一多项式只有16项，分别为

$$x^0y^0+x^0y^1+x^0y^2+x^0y^3+$$
$$x^1y^0+x^1y^1+x^1y^2+x^1y^3+$$
$$x^2y^0+x^2y^1+x^2y^2+x^2y^3+$$
$$x^3y^0+x^3y^1+x^3y^2+x^3y^3$$

其估计结果如图3所示。

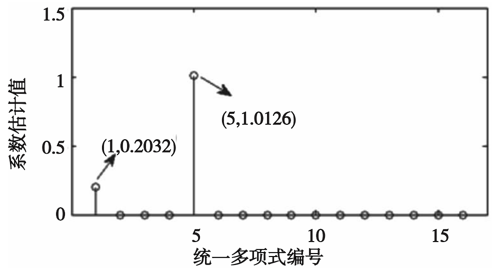

图3　节点N_{12}第二个方程统一多项式系数估计值

图3中的第一项对应x^0y^0即常数项，第五项对应x^1y^0即x，显然图2中也是这么对应的。例如第178项为－100.0472，每个节点都能写成16项，显然178是第12个节点的第2项，即对应x^0y^1。

下面研究噪声对估计误差的影响。对单个节点的估计而言，多项式的非零项或非零项误差有的大于零有的小于零，要比较所有非零项与所有零项的估计误差和时，显然直接用绝对误差会噪声相互抵消，不能充分展示估计误差的真实大小，这里定义噪声的估计误差为绝对误差的绝对值的和，即$E=\sum_i|e_i|$，其中e_i为第i个零项或非零项的绝对误差。图4为第12个节点零项与非零项的估计误差，噪声强度小于0.004时估计误差很小，估计得到的动力学方程和其连接的节点很准确，这里的估计误差可看做误差能量，显然误差能量越小，估计精度越高。其余节点都可得到类似结论。

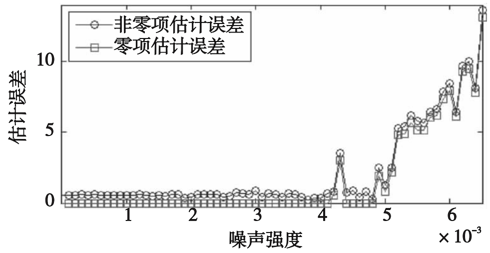

图4　噪声强度对估计值的影响

对每一个节点执行系数估计就得到该网络的拓扑结构，其邻接矩阵如图5所示。图中矩阵外的数值代表节点编号，矩阵中的非零值代表估计得到的两节点间的耦合强度，与图1描述的拓扑结构一致，并且耦合强度非常接近真实值。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
20	0.989	0.965	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.948	0.954	0
19	0.973	0	0	0	0	0	0	0	0	0	0	0.972	0	0	0	0	1.016	0.984	0	0.941
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.923	0.955	0	0.952	0.935
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.962	0.97	0	1.015	0.93	0
16	0.953	0	0	0	0	0	0	0	0	0	0	0	0	0.967	0.984	0	0.989	0.968	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0.996	0.935	0	0.975	0.978	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0.97	0.996	0	0.896	0.956	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0.925	0.982	0	0.962	0.959	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0.994	0.972	0	0.954	0.981	0	0	0	0	0.942	0
11	0	0	0.986	0	0	0	0	0	0.949	0.951	0	0.985	0.981	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0.986	0.97	0	1.021	0.934	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0.969	0.988	0	0.995	0.937	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0.956	1.024	0	0.947	0.974	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0.953	0.956	0	0.989	0.966	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0.973	0.983	0	1.006	0.953	0	0	0	0.954	0	0	0	0	0	0	0	0
5	0	0	0.934	0.981	0	0.96	0.986	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0.965	0.965	0	0.967	0.998	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0.963	0.983	0	0.975	0.955	0	0	0	0	0	0.945	0	0	0	0	0	0	0	0	0
2	0.965	0	0.975	0.97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.021
1	0	0.972	0.967	0	0	0	0	0	0	0	0	0	0	0	0	0.989	0	0	0.997	0.961

图5　FHN小世界网络邻接矩阵估计值

4　结束语

本文应用相关向量机方法实现了小世界神经元网络的节点动力学和拓扑结构估计。以FHN神经元模型为节点、节点间通过膜电压相互电耦合为连边构成的小世界网络为例，通过将耦合方程写成非零项稀疏的统一多项式，利用相关向量机估计得到稀疏解，从而获得节点动力学方程和网络拓扑。结果表明，节点动力学方程结构和网络拓扑均可被准确地估计出来，动力学方程参数和耦合强度也有很高的估计精度，且估计结果对噪声有强鲁棒性。同时，上述结论对网络规模具有鲁棒性，即网络节点数的多少对动力学结构和拓扑估计准确性无影响。本文的研究还表明，该估计策略对Lorenz和Rössler等混沌系统同样有效。

参考文献：

[1] HUANG De-bin. Synchronization-based estimation of all parameters of chaotic systems from time series [J]. Physical Review E, 2004, 69(6): 067201.

(下转第4096页)

值高于 PCSKM 方法，这是因为当数据维数较高时，数据降维过程有效地保持了数据的信息特征，同时忽略了部分冗余信息，提高了聚类的性能。最后，与其他三种方法相比，MFASSC 在五个数据集上取得最高的 RI 值。一方面是因为 MFA 在降维过程中充分利用其类别信息来保持数据的信息特征；另一方面，说明半监督聚类中的监督信息对聚类性能的改善效果显著。

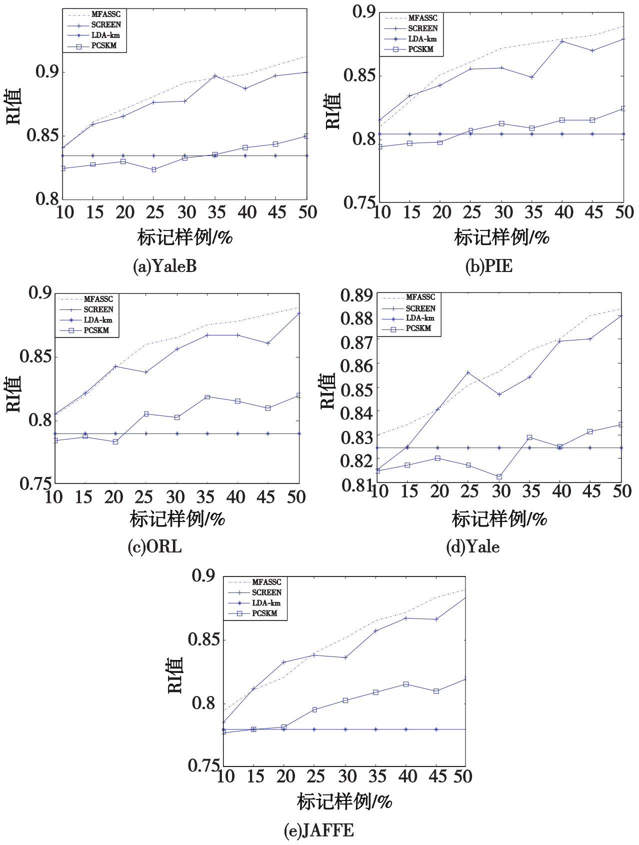

图3　算法在不同数据集上RI值随监督信息量的变化情况

3　结束语

本文提出一种基于间隔 Fisher 分析的半监督聚类算法，算法使用 MFA 对样本进行半监督降维，然后在低维空间中对数据进行半监督聚类，再利用聚类结果指导降维，重复执行降维和聚类过程，直到收敛为止。实验结果表明，该方法能够有效处理高维数据，同时充分利用有监督信息，有效地提高了聚类性能。

在半监督聚类算法中，监督信息越多越有助于提高算法的性能，但获得监督信息的代价较大。如何选择有利的监督信息来提高算法性能成为接下来的研究方向。

参考文献：

[1] BASU S, BANERJEE A, MOONEY R J. Semi-supervised clustering by seeding[C]//Proc of the 19th International Conference on Machine Learning. 2002:19-26.

[2] WAGSTAFF K, CARDIE C, ROGER S, *et al*. Constrained-K-means clustering with background knowledge[C]//Proc of the 18th International Conference on Machine Learning. 2001:577-584.

[3] YEUNG D Y, CHANG Hong. Extending the relevant component analysis algorithm for metric learning using both positive and negative equivalence constraints [J]. Pattern Recognition, 2006, 39(5): 1007-1010.

[4] 王玲，薄列峰，焦李成. 密度敏感的半监督谱聚类[J]. 软件学报，2007,18(10):2412-2422.

[5] CHEN Wei-fu, FENG Guo-can. Spectral clustering: a semi-supervised approach[J]. Neurocomputing,2011,77(1):229-242.

[6] BILENKO M, BASU S, MOONEY R J. Integrating constraints and metric learning in semi-supervised clustering[C]// Proc of the 21st International Conference on Machine Learning. 2004:81-88.

[7] KULIS B, BASU S, MOONEY R J. Semi-supervised graph clustering: a kernal approach[C]//Proc of the 22nd International Conference on Machine Learning. 2005:457-464.

[8] YAN Bo-jun, DOMENICONI C. An adaptive kernal method for semi-supervised clustering[C]// Proc of the 17th European Conference on Machine Learning. 2006:521-532.

[9] TANG Wei, XIONG Hui, ZHONG Shi, *et al*. Enhancing semi-supervised clustering: a feature projection perspective[C]//Proc of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2007:707-716.

[10] DING C H, LI Tao. Adaptive dimension reduction using discriminant analysis and K-means clustering[C]// Proc of the 24th International Conference on Machine Learning. 2007:521-528.

[11] YAN Shui-cheng, XU Dong, ZHANG Ben-yu, *et al*. Graph embedding and extensions: a general framework for dimensionality reduction [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2007,29(1):40-51.

[12] 尹学松，胡恩良，陈松灿. 基于成对约束的判别型半监督聚类分析[J]. 软件学报，2008,19(11),2791-2802.

（上接第 4084 页）

[2] 周芳龙，王浩，姚宏亮. 基于粒子滤波的非线性系统静态参数估计方法[J]. 计算机应用研究，2011,28(5):1637-1639,1654.

[3] YU Dong-chuan, RIGHERO M, KOCAREV L. Estimating topology of networks[J]. Physical Review Letters,2006,97(18):188701.

[4] GOUESBET G, LETELLIER C. Global vector-field reconstruction by using a multivariate polynomial L2 approximation on nets[J]. Physical Review E,1994,49(6):4955-4971.

[5] LU Jun-an, LV Jin-hu, XIE Jin. Reconstruction of the lorenz and chen systems with noisy observations[J]. Computers and Mathematics with Applications,2003,46(8-9):1427-1434.

[6] BEZRUCHKO B P, SMIRNOV D A. Constructing nonautonomous differential equations from experimental time series[J]. Physical Review E,2000,63(1):016207.

[7] WANG Wen-xu, YANG Rui, LAI Ying-cheng, *et al*. Predicting catastrophes in nonlinear dynamical systems by compressive sensing[J]. Physical Review Letters,2011,106(15):154101.

[8] WANG Wen-xu, YANG Rui, LAI Ying-cheng, *et al*. Time-series-based prediction of complex oscillator networks via compressive sensing[J]. Europhysics Letters,2011,94(4):48006.

[9] TIPPING M E. Sparse Bayesian learning and the relevance vector machine[J]. Journal of Machine Learning Research,2001,1:211-244.

[10] NEWMAN M E J, WATTS D J. Renormalization group analysis of the small world network model[J]. Physics Letters A,1999,263(4-6):341-346.