

类内子流形局部间隔对齐的人脸图像判别方法

蒲晓蓉, 樊科, 黄东

(电子科技大学计算机科学与工程学院 成都 611731)

【摘要】提出利用类内子流形对高维人脸图像进行判别分析的新方法,沿对齐的类间局部间隔法向扩展类内子图,获得一系列线性投影,并正交化线性投影得到一组子空间的正交基向量。使用局部相邻关系增大类间差异,并将类内结构保存到与类间间隔区域对齐的子空间中,能有效降低因人脸图像拍摄角度、姿态、光照、眼镜和性别等因素导致的数据多模态或低维流形的高曲率对识别性能的影响。在Yale Face Database B和UMIST face database上进行的实验证明,较之LPP和FDA等方法,该方法能更加可靠地保留类内和类间的子流形结构,且有更高的识别准确率。

关键词 判别分析; 人脸识别; 特征提取; 图像分类

中图分类号 TP39

文献标识码 A

doi:10.3969/j.issn.1001-0548.2010.06.023

Discriminant Analysis of Face Images by Local Margin Alignment

PU Xiao-rong, FAN Ke, and HUANG Dong

(School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 611731)

Abstract This paper proposes a new approach to perform the discriminant analysis on the labelled high dimensional image data with intra-class sub-manifolds. Real world images are usually taken from the different camera views. Pose, illumination, glasses and gender of the persons taking the facial images usually lead to multi-modality or high curvature of the underlying manifold structures. These variations result in the degraded performance of many existing algorithms. This paper proposes to preserve the within-class local structure, while imposing constrain on the variances only in the directions normal to the between-class margin. The experiments on Yale-B and UMIST face database show that the proposed algorithm outperforms many approaches such as LPP (locality preserving projections) and FDA (fisher discriminant analysis).

Key words discriminant analysis; face recognition; feature extraction; image classification

人脸图像数据在原输入空间中具有极高的维度(常高于 10^4)。尽管如此,图像数据中通常存在隐含的非常低维的子流形^[1]。将局部信息用于高维图像数据的判别分析是目前非常重要的研究方向^[2]。

传统的分析方法以主元分析(PCA)^[3]和Fisher线性判别分析(FDA)^[4]为代表。由于它们只关心全局的欧氏子空间结构,因此处理多模态或者非高斯分布的数据比较困难。考虑局部结构的流形学习^[1]是处理上述问题的一种有效方法。流形方法可以分为线性和非线性投影方法。本文主要关注在人脸识别中适用的线性方法,该类方法可以得到明确的投影矩阵,便于计算新样本在特征空间的投影。

文献[5]提出局部保持投影(LPP)对拉普拉斯特征映射进行线性逼近。LPP用于人脸识别(laplacianface)^[6]获得了优于PCA和LDA的效果。局部Fisher判别分析(LFDA)^[7]有效地结合了FDA和

LPP的思想,通过局部变尺度方法(local scaling method)^[8]构建带权重的连通图,将类内结构分成多模态簇,而其最优投影通过同时最小化局部类内差异和最大化剩下的差异得到。间隔Fisher判别分析(MFA)^[9]将LPP扩展到类内和类间子图。MFA构建一个本征图和一个惩罚图分别刻画类内的紧密性和类间的可分性。LFDA和MFA都得到了较好的识别准确率,但LFDA在子流形呈多模态的情况下非常有效,MFA更适用于子流形连通且非凸的情况。

本文提出一种新的判别方法——局部间隔对齐(LMA)。LMA的类内表示基于LPP方法;类间间隔的结构描述为一种局部化的、类似FDA的类间方差。LMA方法能沿着对齐的类间局部间隔法向扩展类内子图,并获得一系列线性投影,通过正交化线性投影能得到一组子空间的正交基向量。最终,LMA的投影可以表示为一个保留原空间尺度结构的旋转

收稿日期: 2009-04-15; 修回日期: 2010-03-28

基金项目: 国家自然科学基金(60973070); 四川省产业技术与开发项目

作者简介: 蒲晓蓉(1969-),女,博士,主要从事生物特征识别、计算智能等方面的研究。

映射。LMA参考了许多已存在算法的结构保持特性,使其对多模态和非凸的子流形的判别分析都有较好的效果。

1 局部间隔对齐算法

假设有 c 类数据,第 i 类的个数为 n_i ($i=1,2,\dots,c$)且 $\sum_i n_i = n$ 。输入数据 $X_i = \{x_j \in \mathbf{R}^d\}$ ($j=1,2,\dots,n_i$)是第 i 类中可能含噪音的采样,如图1a所示。设数据是局部线性可分的,通过本文方法将输入数据 $X \in \mathbf{R}^{d \times n}$ 降维得到低维特征向量 $Z = [z_1, z_2, \dots, z_n] \in \mathbf{R}^{m \times n}$ ($m \ll d$)进行判别分析。类间间隔区域和类内结构如图1b和图1c所示。

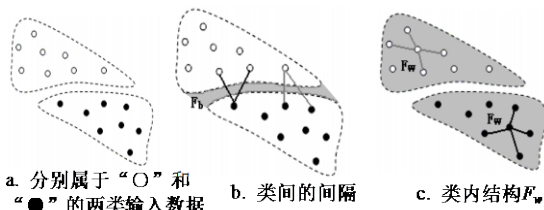


图-1 两类问题的几何结构示意图。

假设光滑的间隔区域 F_b 位于 c 类 d 维输入数据间,如图1b所示。对第 i 类数据,希望通过间隔区域获取所有类之间的差异,找到局部的方向向量能将其与其他 $c-1$ 类数据尽量分开。注意到所有的类间邻居都处于 $c-1$ 类的靠近第 i 类的边界上,因此所有的间隔区域 F_b 可以看作一个横跨这些局部方向的子空间。通过对齐这些局部方向可以得到该子空间,就能最大分离所有的 c 类数据。

另外,真实世界的图像和视频数据中,类内数据也有自身的低维结构 F_w ,如图1c所示。例如,人脸图像不同的姿态、光线和表情的变化都能通过基于类内邻接图的流形学习算法学习。因此允许预测新的输入数据在特征空间的位置,从而进一步提高投影的泛化能力。本文算法力求在输入数据投影到特征空间时保留类内的局部几何结构,完全不同于采用Fisher判别式最小化类内变化的传统方式。本文算法的主要思想为:(1)通过对齐类间局部间隔来最大化类间的可分性;(2)通过对齐类内结构保证类内的泛化能力。

1.1 类间局部间隔

首先,对数据点 x_t ($t=1,2,\dots,n$)进行局部间隔建模。不失一般性,考虑 k 局部线性可分的情况。

定义 如果存在线性分类器将数据点 x 与 k 个最近的类间邻居分开,则称 x 是 k 局部线性可分的。

k 局部线性可分的几何解释如图2所示。可以在

点 x_t 及其 k_b 个类间最近邻居 x_j ($j \in N_{k_b}(t)$)之间求解一个最优的分界面。其中, $N_{k_b}(t)$ 表示 x_t 的 k_b 个根据欧式距离排序的类间邻居的集合,该分界面的方向称为局部判别方向。通过求解优化问题
$$\max_{\mathbf{P}^{(t)}} \left\| \left(\mathbf{P}^{(t)} \right)^T \sum_{j \in N_{k_b}(t)} \alpha_j^{(t)} (x_t - x_j) \right\|_2^2$$
得到最优的方向向量 $\mathbf{P}^{(t)} \in \mathbf{R}^d$,其中 $\alpha_j^{(t)}$ ($j \in N_{k_b}(t)$)是相应最近类间邻居的权重。注意, $x_t \rightarrow x_j$ 是有向的,也就是说,若存在 $j \in N_{k_b}(t)$,但不一定存在 $t \in N_{k_b}(j)$ 。因此,类间间隔被局部化,如图2所示。

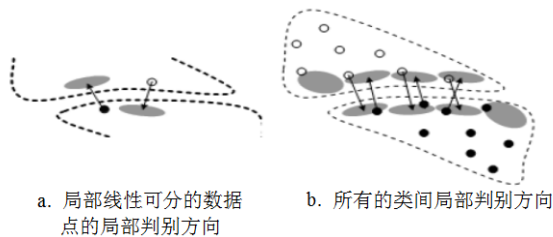


图-2 局部线性可分的几何解释

对于权重 $\alpha_j^{(t)}$,本文提出两种计算方式。

(1) 简化方法:若 $j \in N_{k_b}(t)$,则 $\alpha_j^{(t)} = 1$,否则 $\alpha_j^{(t)} = 0$,则最优的局部判别方向:

$$\mathbf{p}^{(t)} = \text{Normalize} \left\{ x_t - \frac{1}{k_b} \sum_{j \in N_{k_b}(t)} x_j \right\}$$

(2) 平滑局部尺度方法⁰:

$$\alpha_j^{(t)} = \exp \left(- \frac{\|x_t - x_j\|_2^2}{\sigma^2} \right) \quad j \in N_{k_b}(t)$$

式中 σ^2 是局部尺度参数。在没有可行的优先选择的情况下,可以估计:

$$\sigma^2 = \|x_t - x_{k_b}^{(t)}\|_2 \cdot \|x_{k_b}^{(t)} - (x_{k_b}^{(t)})^{(t)}\|_2$$

式中 $x_{k_b}^{(t)}$ 是 x_t 的第 k_b 个最近类间邻居。此时有

$$\mathbf{p}^{(t)} = \text{Normalize} \left\{ \frac{\sum_{j \in N_{k_b}(t)} \alpha_j^{(t)} (x_t - x_j^{(t)})}{\sum_{j \in N_{k_b}(t)} \alpha_j^{(t)}} \right\}$$

由此,通过局部线性可分性可以获得线性变换 $z_t = (\mathbf{P}^{(t)})^T x_t$ 。最后,全局优化地将所有点的局部判别方向进行对齐,得到一个全局映射 $\mathbf{X} \rightarrow \mathbf{Z}$:
$$\mathbf{Z} = (\mathbf{P}_{k_b})^T \cdot \mathbf{X} \quad (\mathbf{P}_{k_b} \in \mathbf{R}^{d \times m})$$
。于是,类间的可分性最大化为:

$$\max_Z \sum_{t=1}^n W_{k_b}^{(t)} \left\| \sum_{j \in N_{k_b}(t)} \alpha_j^{(t)} (x_t - x_j^{(t)}) \right\|_2^2 =$$

$$\max_Z \sum_{t=1}^n W_{k_b}^{(t)} \left\| \sum_{j \in N_{k_b}(t)} \alpha_j^{(t)} (\mathbf{P}_{k_b})^T (x_t - x_j^{(t)}) \right\|_2^2 =$$

$$\max_Z \text{Trace}[(\mathbf{P}_{k_b})^T \mathbf{X} (\mathbf{I}_n - \mathbf{A}_{k_b})^T \mathbf{W}_{k_b} (\mathbf{I}_n - \mathbf{A}_{k_b}) \mathbf{X}^T \mathbf{P}_{k_b}]$$

式中 $\mathbf{I}_n \in \mathbf{R}^{n \times n}$ 是单位阵; $\mathbf{A}_{k_b} \in \mathbf{R}^{n \times n}$ 是一个系数矩

$$\text{阵, } A_{k_b}(t, j) = \frac{\alpha_j^{(t)}}{\sum_{j=1}^n \alpha_j^{(t)}} \quad (t, j=1, 2, \dots, n); \mathbf{W}_{k_b} \in \mathbf{R}^{n \times n}$$

是一个对角阵, 其对角元 $W_{k_b}^{(t)} = \frac{\sum_{j=1}^n \alpha_j^{(t)}}{\sum_{\{t|x_t \in X_i\}} \sum_{j=1}^n \alpha_j^{(t)}}$

是第 i 类中 $x_t (\{t|x_t \in X_i\})$ 的归一化的权重因子。本文中, $W_{k_b}^{(t)}$ 越大, x_t 在决定全局判别方向时越重要。

1.2 类内局部结构

与类间间隔相似, x_t 点附近的类内结构可以用其与其 k_w 个类内最近邻居 x_j 间的加权边表示为 $j \in N_{k_w}(t)$, 其中 $N_{k_w}(t)$ 表示 x_t 的 k_w 个根据欧式距离排序的类内邻居的集合。本文希望, 将输入空间的加权图投影到特征空间时, 能够使子流形上连通的点始终靠近。借鉴 LPP^[5] 的思想, 通过优化

$$\min_Z \sum_{t=1}^n W_{k_w}^{(t)} \|z_t - z_j\|_2^2$$

的结果保留类内结构, 其中 $\mathbf{Z} = [z_1, z_2, \dots, z_n] \in \mathbf{R}^{m \times n}$ 由特征空间的向量组成; $W_{k_w}^{(t)}$ 是类内边 $x_t \leftrightarrow x_j$ 的权重因子, $j \in N_{k_w}(t)$ 或 $t \in N_{k_w}(j)$ 。与类间间隔不同, 类内边是无向的, 所以权重矩阵 $\mathbf{W}_{k_w} \in \mathbf{R}^{n \times n}$ 是对称阵, 即 $W_{k_w}^{(t)} = W_{k_w}^{(j)}$ ($t, j=1, 2, \dots, n$)。设 $z_t = (\mathbf{P}_{k_w}^{(t)})^T x_t$, 其中 $\mathbf{P}_{k_w} \in \mathbf{R}^{d \times m}$, 则上述的优化问题可以改写为:

$$\min_Z \sum_{t=1}^n W_{k_w}^{(t)} \left\| (\mathbf{P}_{k_w}^{(t)})^T (x_t - x_j) \right\|_2^2 =$$

$$\min_Z \text{Trace}[(\mathbf{P}_{k_w}^{(t)})^T \mathbf{X} (\mathbf{D}_{k_w} - \mathbf{W}_{k_w}) \mathbf{X}^T \mathbf{P}_{k_w}^{(t)}]$$

式中 \mathbf{D}_{k_w} 是对角阵, 其对角元 $(\mathbf{D}_{k_w})_{tt} = \sum_j W_{k_w}^{(t)}$ 。

同样, 有两种确定权重的方法:

(1) 简化方法: $W_{k_w}^{(t)} = 1$ 。

(2) 平滑局部尺度方法: $W_{k_w}^{(t)} = \exp\left(-\frac{\|x_t - x_j^{(t)}\|_2^2}{\sigma^2}\right)$,

其中 $\sigma^2 = \|x_t - x_{k_w}^{(t)}\|_2 \cdot \|x_{k_w}^{(t)} - (x_{k_w}^{(t)})_{k_w}^{(t)}\|_2$ 。

对点 x_t 而言, x_j 越靠近它, 影响边 $\{x_t \leftrightarrow x_j^{(t)}\}$ 的 $W_{k_w}^{(t)}$ 就越大。换句话说, $(\mathbf{D}_{k_w})_{tt}$ 越大, x_t 在类内结构中越重要。由此可见, 上述的优化问题就是试图确

保在输入空间相互接近的点, 在特征空间也保持邻近。由于权重因子 $W_{k_w}^{(t)}$ 是通过每个局部区域计算获得的, 从而保持了局部关系。

1.3 线性局部间隔对齐

非线性映射可以使高维数据的潜在几何结构在可视化的低维子空间显示出来。但是, 在实际使用中, 尤其是聚类 and 图像识别等情况下, 需要在输入新数据进行样本外扩展时进行全局邻居搜索。而线性投影由于具有闭合解析解, 计算复杂度大大降低, 所以具有相当的应用优势。

假设 $\mathbf{X} \in \mathbf{R}^{d \times N}$ 和 $\mathbf{Z} \in \mathbf{R}^{m \times N}$ 之间的一对一的映射可以定义为 $\mathbf{Z} = \mathbf{P}^T \mathbf{X}$, 其中 $\mathbf{P} \in \mathbf{R}^{d \times m}$, 则上述全部优化可重写为:

$$\min_P \mathbf{P}^T \mathbf{S}_w \mathbf{P} \quad \text{s.t.} \quad \mathbf{P}^T \mathbf{S}_b \mathbf{P} = \mathbf{I}_m$$

式中 $\mathbf{I}_m \in \mathbf{R}^{m \times m}$ 是单位阵, 且 $\mathbf{S}_w = \mathbf{X} (\mathbf{D}_{k_w} - \mathbf{W}_{k_w}) \mathbf{X}^T$, $\mathbf{S}_b = \mathbf{X} (\mathbf{I}_n - \mathbf{A}_{k_b})^T \mathbf{W}_{k_b} (\mathbf{I}_n - \mathbf{A}_{k_b}) \mathbf{X}^T$ 。该问题又可以转化为广义特征值问题 $\mathbf{S}_w \mathbf{P} = \mathbf{S}_b \mathbf{P} \mathbf{\Lambda}$ 取最小的 m 个特征向量求解, 其中 $\mathbf{\Lambda}$ 为广义特征值对角矩阵。

经过投影, 类内邻近的一对数据在特征空间也是相互靠近的, 而不同类的数据则相互分离。另外, 本文算法还具有相隔较远同类数据不会被强迫靠近在一起的特性。

广义特征向量 $\mathbf{P} \in \mathbf{R}^{d \times m}$ 确定了一个判别子空间, 与分离性没有冲突的类内结构都能得以保留。

由于正交基能够提供统一的局部度量:

$$\text{dist}(z_i, z_j) = \|z_i - z_j\|_2 = \left\| \mathbf{P}^T (x_i - x_j) \right\|_2 = \sqrt{(x_i - x_j)^T \mathbf{P} \mathbf{P}^T (x_i - x_j)}$$

如果 $\mathbf{P} \mathbf{P}^T = \mathbf{I}$, 则特征空间的距离 $\text{dist}(z_i, z_j)$ 近似于其在输入空间的距离 $(x_i - x_j)^T (x_i - x_j)$ 。为了保留尺度信息, 本文采用施密特正交化将 \mathbf{P} 转换为一组正交基得到最终的 LMA 线性投影矩阵。

2 实验与讨论

为了验证本文局部间隔对齐算法的有效性, 使用如图3和图4所示的 Yale Face Database B^[10] 和 UMIST face database^[11] 数据库与 LPP、FDA、LFDA 和 MFA 算法进行对比。



图3 Yale B数据库中的两个被试的部分样本



图4 UMIST中的某被试的部分样本

由于输入的高数据维数带来的矩阵病态, 实验数据统一使用线性PCA进行预处理(选取特征值总和的前99.99%)。另外, 由于局部方法的性能和选取的邻居个数有关, 在实验中使用统一的邻居数 k (简化起见MFA和LMA中 $k_w = k_b = k$)。每一个图片集被分成测试集(R)和训练集(E)两部分。以“R4Ek3”为例, 表示对每一个人随机地选取4张训练图片, 所有没有用于训练的图片作为测试集, 邻居数 $k=3$ 。分类算法使用 k 近邻分类器。

2.1 Yale B人脸数据集

该数据库包含10个对象的各576张图像, 共计9种姿态 \times 64种光照条件的5760张灰度图片, 图像尺寸被统一为 100×100 像素点。其中姿态的不连续变化使得数据潜在的子流形呈多簇, 而光照的连续变化使得数据分布连通非凸。所有算法选取不同的随机训练集20次, 平均识别准确度(%)比较结果见表1和表2。表1中的类根据对象标记, 随机选择10人每人576张图片中的1000张(100张/人)作为训练集, 剩余的4760张作为测试集。表2中的类根据姿态标记, 随机选择的 9×100 张图片作为训练集, 剩余的4860张作为测试集。表1和表2的实验结果表明LMA的性能优于所对比的其他算法。

表1 Yale B上识别被试身份的比较结果

算法	LPP	FDA	LFDA	MFA	LMA
R100Ek5	89.0236	90.46,9	88.88,40	80.34,40	91.78,24
R100Ek10	88.3740	90.01,9	87.60,40	81.52,40	91.50,25
R100Ek20	88.1940	89.96,9	82.90,40	82.90,40	91.31,27

表2 yaleB上识别被试姿态的比较结果

算法	LPP	FDA	LFDA	MFA	LMA
R100Ek5	94.11,38	85.43,8	85.51,16	75.84,40	86.70,13
R100Ek10	83.82,38	85.29,8	85.2521	77.03,40	87.06,15
R100Ek20	83.98,40	85.37,8	85.14,25	79.42,40	87.47,27

2.2 UMIST人脸数据集

该数据集包含20个对象共564张人脸图像, 每个对象都包含从正面到侧面的连续姿态变化, 类内结构随着训练集的增加由多模态变为连通的单模态。所有算法选取不同的随机训练集50次, 平均识别准确度(%)比较结果如表3所示。随着训练集的增加, FDA和LPP的准确率与LFDA、MFA和LMA逐步接

近。这是由于训练集越大, 类内局部结构越趋于连通非凸而非多模态, 于是LPP、LFDA、MFA和LMA的局部保持作用被减弱。另外, LFDA和MFA在投影维度较小时都有稳定的识别率。LFDA在训练集稀疏时表现优秀, 而MFA更适合训练集稠密的情况。从表3可见LMA的准确率和稳定性优于所对比的算法。

表3 UMIST数据集上相关算法比较结果

算法	LPP	FDA	LFDA	MFA	LMA
R6Ek5	86.94,18	87.63,19	93.75,17	92.05,40	94.97,30
R8Ek7	90.90,18	91.52,19	95.70,14	95.00,36	97.23,36
R10Ek9	93.49,18	94.12,19	95.39,13	96.09,39	97.35,39
R12Ek11	95.41,18	95.81,19	95.38,10	97.35,39	98.36,40
R12Ek4	95.34,24	95.75,19	95.68,10	97.12,40	98.91,40
R14Ek4	96.78,19	97.11,19	93.86,8	97.87,40	99.20,36

3 结论

本文提出了一种将局部信息结合进线性判别分析的局部间隔对齐算法, 该方法使用局部的相邻关系来增大类间差异, 并且类内结构被保存到与类间隔区域对齐的子空间中, 使得该方法比LPP、FDA、LFDA和MFA更可靠地保留了类内和类间的子流形结构。在人脸数据集上的实验证明, 本文方法比相关算法有更高的识别准确率。下一步将改进算法来以解决欠采样数据或不完整的数据分类问题, 拟引入自适应的邻居选择策略。

本文研究工作得到电子科技大学青年科技基金(L08010601JX0674)的资助, 在此表示感谢。

参考文献

- [1] TENENBAUM J B, SILVA V D, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290: 2319-2323.
- [2] 蒲晓蓉, 郑自明, 周伟. 基于子空间投影的残缺人脸图像识别新算法[J]. 电子科技大学学报, 2006, 35(2): 208-220.
- [3] TURK M, PENTLAND A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.
- [4] BELHUMEUR P N, HEPANHA J P, KRIEGMAN D J. Eigenfaces vs fisherfaces: Recognition using class specific linear projection[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [5] HE X, NIYOGI P. Locality preserving projections[C]// Proceedings of Neural Information Processing System. Vancouver, Canada: MIT Press, 2003: 153-160.
- [6] HE X, YAN S, HU Y, et al. Face recognition using laplacianfaces[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2005, 27(3): 328-340.
- [7] SUGIYAMA M. Dimensionality reduction of multimodal

- labeled data by local Fisher discriminant analysis[J]. Journal of Machine Learning Research, 2007, 8: 1027-1061.
- [8] ZELNIK M L, PERONA P. Self-tuning spectral clustering [C]//Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2005: 1601-1608.
- [9] YAN S, XU D, ZHANG B, et al. Graph embedding and extensions: a general framework for dimensionality reduction[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2007, 29(1): 40-51.
- [10] GEORGHIADES A S, BELHUMEUR P N, KRIEGMAN D J. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2001, 3(6): 643-660.
- [11] GRAHAM D B, ALLINSON N M. Face recognition: from theory to applications[M]. [S.l.]: Springer, 1998: 446-456.

编辑 蒋 晓

(上接第914页)

参 考 文 献

- [1] PEISERT S, BISHOP M, KARIN S, et al. Analysis of computer intrusions using sequences of function calls[J]. IEEE Trans on Dependable and Secure Computing, 2007, 4(2): 137-150.
- [2] MA Xin-xin, ZHAO Yang, QIN Zhi-guang. Improving resilience against DDoS attack in unstructured P2P networks[J]. Journal of Electronic Science and Technology of China, 2007, 5(1): 18-22.
- [3] HERRERIAS J, GOMEZ R. A log correlation model to support the evidence search process in a forensic investigation[C]// Proceedings of the Second International Workshop on Systematic Approaches to Digital Forensic Engineering. New York: IEEE Computer Society Press, 2007: 31-42.
- [4] 孙 波. 计算机取证方法关键问题研究[D]. 北京: 中国科学院软件研究所, 2004.
SUN Bo. Research on key aspects of computer forensic methods[D]. Beijing: Institute of Software of Chinese Academy of Sciences, 2004.
- [5] ABRAHAM T, VEL O. Investigative profiling with computer forensic log data and association rules[C]// Proceedings of the 2002 IEEE International Conference on Data Mining. New York: IEEE Computer Society Press, 2002: 11-18.
- [6] ABRAHAM T. Event sequence mining to develop profiles for computer forensic investigation purposes[C]//Proceedings of the 2006 Australasian workshops on Grid computing and e-research. Darlinghurst, Australia: Australian Computer Society, 2006: 145-153.
- [7] CASTRO L N, ZUBEN F J. The clonal selection algorithm with engineering applications[C]//Proceedings of GECCO'00, Workshop on Artificial Immune Systems and Their Applications. New York: ACM Press, 2000: 36-37.
- [8] TIMMIS J, HONE A, STIBOR T, et al. Theoretical advances in artificial immune systems[J]. Theoretical Computer Science, 2008, 403(1): 11-32.
- [9] KHILWANI N, PRAKASH A, SHANKAR R, et al. Fast clonal algorithm[J]. Engineering Applications of Artificial Intelligence, 2008, 21(1): 106-128.
- [10] 焦李成, 刘 芳, 刘 静, 等. 智能数据挖掘与知识发现[M]. 西安: 西安电子科技大学出版社, 2006.
JIAO Li-cheng, LIU Fang, LIU Jing, et al. Intelligent data mining and knowledge discovery[M]. Xi'an: Xidian University Press, 2006.
- [11] 金可仲. 基于关键属性约束的关联规则挖掘在日志分析中的应用[J]. 温州大学学报, 2008, 29(1): 56-60.
JIN Ke-zhong. Application on log analysis using association rule mining algorithm with key item constraint[J]. Journal of Wenzhou University, 2008, 29(1): 56-60.
- [12] AGRAWAL R, SRIKANT R. Fast algorithm for mining association rules[C]//Proceedings of the 20th Very Large Data Bases International Conference. San Francisco: Morgan Kaufmann Publishers, 1994: 487-499.
- [13] 杨 珺, 王 敏, 陈 晨, 等. 带约束的免疫克隆取证分析方法[J]. 计算机工程, 2010, 36(14): 158-160.
YANG Jun, WANG Min, CHEN Chen, et al. Immune clone forensic analysis method with constraint[J]. Computer Engineering, 2010, 36(14): 158-160.
- [14] PAXSON V. Traces available in the internet traffic archive[EB/OL]. (2008-04-09) [2008-07-30]. <http://ita.ee.lbl.gov/htm/traces.html>.

编辑 张 俊