

犯罪预警的乘积积分模型研究*

王琦, 曹圣山**, 陈丕炜, 闫盼盼, 吕可波, 高翔

(中国海洋大学数学科学学院, 山东 青岛 266100)

摘要: 本文通过建立数学模型来对犯罪嫌疑人进行分类分级。针对相关职能部门已投入使用的犯罪预警的积分模型中存在的问题, 本文基于某市犯罪嫌疑人的警情数据信息进行分析处理, 提取特征指标并选择合适的量化方式将指标量化后, 构建了犯罪预警的乘积积分模型, 并将乘积积分模型参数的求解归结为一个参数拟合问题。针对参数拟合问题中目标函数非解析的特点, 运用 Monte Carlo 方法进行求解从而确定近似最优参数。根据构建的模型计算犯罪嫌疑人的积分, 确定其预警级别的高低, 进而对犯罪嫌疑人分类分级。数值实验验证了本文构建的乘积积分模型与现有积分模型相比具有较好的准确率, 为提高侦查效率奠定了基础。

关键词: 犯罪预警; 乘积积分模型; 分类分级; 参数拟合; Monte Carlo 方法

中图分类号: TP391

文献标志码: A

文章编号: 1672-5174(2016)10 [I]-154-05

DOI: 10.16441/j.cnki.hdx.20150383

引用格式: 王琦, 曹圣山, 陈丕炜, 等. 犯罪预警的乘积积分模型研究[J]. 中国海洋大学学报(自然科学版), 2016, 46(增): 154-158.

WANG Qi, CAO Sheng-Shan, CHEN Pi-Wei, et al. Research on product integral model of early warning[J]. Periodical of Ocean University of China, 2016, 46(Sup.): 154-158.

近年来, 犯罪手段日趋多样, 犯罪形势复杂, 增加了公共安全保障及社会管理的难度和成本^[1]。随着信息化的推进, 国内许多地方政府的公共安全部门已经收集了大量的犯罪嫌疑人的警情信息并建立了相应的信息系统^[2], 为相关部门提供了强大的综合分析研判工具, 也为建立数学模型奠定了良好的基础^[3]。通过对警情数据进行分析 and 判断不仅可以发现蕴藏于数据背后的一些重要规律, 揭示社会安全状态的变化特征, 还能够为未来某一阶段可能出现的警情情况进行预测, 使相关部门对可能发生的突发事件做好应急准备^[4]。因此, 通过合理的警情研判和预测分析, 实现内外部资源的整合共享和综合利用, 对犯罪嫌疑人进行分类分级, 进而作出不同的侦查管理决策成为相关部门重点关注的问题。在警力不足的情况下, 有助于掌握主动权、先发打击犯罪。

Mena J. 介绍了数据挖掘技术如智能代理、文本挖掘、神经网络、机器学习等方法有助于安全与犯罪侦查, 阐述了如何构建预测模型来检测犯罪活动并对犯罪嫌疑人进行行为分析^[5]; BKL Fei, JHP Eloff 运用自组织映射检测异常行为^[6]; 国内也有相关的研究, Wang H, R Hu 等分别基于决策树分析犯罪行为^[7-8]。

而目前国内针对犯罪嫌疑人管控主要投入使用的则是积分模型。如北京、西藏、云南等地已分别建立并使用犯罪预警的积分模型。这些积分模型构建的方式基本相同。首先, 选择特征指标集; 其次, 根据特征指标含义, 确定嫌疑人不同表现的积分值; 最后, 对各嫌疑人将特征指标集中所有指标的积分值直接相加得到综合积分值, 并由此积分值的高低决定嫌疑人的预警级别高低。模型结构形如 $G_i = \sum_j g_{ij}$, 其中 G_i 为第 i 个嫌疑人的综合积分, g_{ij} 为第 i 个嫌疑人第 j 个指标的指标表现的积分值。现有的积分模型主要存在两个问题: 其一, 将指标的重要性与嫌疑人的指标表现合二为一来确定积分值; 其二, 特征指标积分值的确定依靠主观经验。积分模型的本质是确定嫌疑人的危险程度的高低, 现有积分模型的积分仅由个人的特征指标表现确定, 缺乏群体行为间的相互影响度量以及对指标重要性的刻画。

针对相关职能部门已投入使用的积分模型中存在的问题, 本文基于某市犯罪嫌疑人的警情信息, 研究构建犯罪预警的乘积积分模型, 并基于 Monte Carlo 方法进行参数拟合确定模型的近似最优参数, 数值实验验证了方法的有效性, 与现有的积分模型相比, 提高了模

* 基金项目: 国家自然科学基金项目(11071228)“关于图像处理模型的目标函数构造及其数值方法研究资助”

Supported by NNSF of China(11071228)“Study on the Structure and Numerical Methods of Object Function in Image Processing Model”

收稿日期: 2015-04-15; 修订日期: 2015-11-12

作者简介: 王琦(1991-), 女, 硕士, 主研方向: 数学建模及其数值解法研究。E-mail: Shaqizi315923@163.com

** 通讯作者: E-mail: sscao@ouc.edu.cn

型准确率。

1 特征提取及数据处理

1.1 数据来源

本研究和某市司法机关合作,并得到其提供的研究数据,包含 2004—2014 年犯罪嫌疑人的各类警情信息。嫌疑人数据库中数据信息的不同来源使得原始数据取值方式具有多样性,包括数值、文字、日期、字典项等多种格式,部分数据内容如表 1:

表 1 犯罪嫌疑人信息表

Table 1 Original information of criminal suspects

编号 Serial Number	性别 Sex	民族 Nation	年龄段 Age group	文化程度 Educational level
00001	男	汉	36~45	高中
00002	男	汉	36~45	初中
00003	女	汉	26~35	初中
.....

1.2 特征提取

模型所需的特征指标的提取基于以下三个方面,其一是犯罪行为学和犯罪心理学的基本原理^[9];其二是专业人员的实践经验;其三是根据犯罪嫌疑人数据信息特征将数据缺失严重、数据异常、覆盖率不足 1/3 的某些指标予以删除。据此,提取了犯罪嫌疑人包括性别、年龄等基本信息以及行为信息在内的 N 个特征指标的数据信息。

1.3 数据处理

对任意(有限)集合 A ,用 $|A|$ 表示 A 所含元素个数。 M 表示在库犯罪嫌疑人集合,记 $m = |M|$,并将 M 中的元素标号为 $(1, 2, \dots, m)$;用 $X = (x_{ij})_{m \times N}$ 表示在库犯罪嫌疑人的原始数据。

根据各个指标的数据特征及其实际意义确定其各自的数量化方式,记 $F(X)$ 为指标数量化过程,其中 $F = (F_1, F_2, \dots, F_N)$ 表示指标的数量化方式。模型包括单调量化、正态量化、统计量化、相对统计量化、统计倒数量化 5 种数量化方式。

假设对第 j ($j = 1, 2, \dots, N$) 列属性的数据进行数量化处理, $Y = (y_{ij})_{m \times N}$ 表示数量化后的数据,5 种数量化方式的处理方式分别为:

- (1) 单调量化: $y_{ij} = x_{ij}$ ($i = 1, 2, \dots, m$)。
- (2) 正态量化:对样本值 $x_{1j}, x_{2j}, \dots, x_{mj}$, 求均值 μ_j 和标准差 σ_j ; 计算 $y_{ij} = e^{-\frac{(x_{ij} - \mu_j)^2}{2\sigma_j^2}}$ ($i = 1, 2, \dots, m$)。
- (3) 统计量化:计算样本 $x_{1j}, x_{2j}, \dots, x_{mj}$ 可能的取值 c_1, c_2, \dots, c_s ; 对每个可能取值计算样本值 $x_{1j}, x_{2j}, \dots, x_{mj}$

出现的频数或频率 f_1, f_2, \dots, f_s ; 如果 $x_{ij} = c_v$, 则 $y_{ij} = f_v$ ($v = 1, 2, \dots, s$)。

(4) 相对统计量化:计算样本 $x_{1j}, x_{2j}, \dots, x_{mj}$ 可能的取值 c_1, c_2, \dots, c_s ; 给定相对向量 C_1, C_2, \dots, C_s (C 为 c 的总值); 对每个可能取值计算样本值 $x_{1j}, x_{2j}, \dots, x_{mj}$ 出现的频数或频率 f_1, f_2, \dots, f_s ; 如果 $x_{ij} = c_v$, 则 $y_{ij} = \frac{f_v}{C_v}$ ($v = 1, 2, \dots, s$)。

(5) 倒数统计量化:与统计量化方式基本相同,不同之处在于如果 $x_{ij} = c_v$, 则 $y_{ij} = \frac{1}{f_v}$ ($v = 1, 2, \dots, s$)。

为消除量纲的不同,对数量化后的数据再进行标准化处理,使其具有统一的均值和标准差,用 $Y' = (y'_{ij})_{m \times N}$ 表示标准化后数据且 $y'_{ij} > 0$ 。处理后的部分数据如表 2:

表 2 犯罪嫌疑人数据处理后信息表

Table 2 Information of criminal suspects after data processing

编号 Serial number	性别 Sex	年龄 Age	文化程度 Educational level	户籍地 Registered residence
0000 1	1. 132 3	1. 207 0	0. 459 1	0. 888 0
0000 2	1. 132 3	1. 242 6	1. 119 1	0. 728 9
0000 3	0. 527 8	1. 201 8	1. 119 1	1. 020 6
.....

2 消除特征指标相关性

在评估犯罪嫌疑人危险程度的非线性数学模型结构中,为了使所确定的特征指标既全面又不重复,也为了接下来运用 Monte Carlo 方法进行最优参数拟合^[10],应该注意相关性这一环节的处理,尽可能保证各指标间相互独立。相关性在实际评价中应予高度重视并应尽力消除,否则将会由于指标的重复计算而直接影响评估的有效性和可靠性^[11]。

本文参照文献^[12-13]的方法,将定性分析与定量计算相结合,基于指标相关矩阵,对指标集合进行等价类划分,对同一等价类的指标用一个指标综合函数表示,然后利用随机向量的相关性,进行优化求解,求得等价类指标综合函数,使等价类之间的相关性最小,以消除指标相关性。

本文基于某市在库犯罪嫌疑人数据,由上述分析确定的 N ($N = 37$) 个指标 $U = \{u_1, u_2, \dots, u_N\}$ 及相应的标准化后数据 Y' , 可得其相关矩阵 R , 求 R 的传递闭包 $t(R)$, 选取水平值 $\lambda = 0.8$, 则可得 n 个等价类 $U^{(i)}$ ($i = 1, 2, \dots, n$), 将包含有 2 个指标及以上的等价类找出, 分别为 $U^{(17)} = (u_{17}, u_{19})$ 、 $U^{(24)} = (u_{25}, u_{26})$ 、

$U^{(31)} = (u_{33}, u_{36}, u_{37})$, 计算其协方差阵 \sum , 并优化求解可得 3 个等价类指标综合函数^[12-13]分别为:

$$u(U^{(17)}) = 0.973 2u_{17} + 0.230 1u_{19};$$

$$u(U^{(24)}) = 1.000 0u_{25} + 0.000 0u_{26};$$

$$u(U^{(31)}) = 0.022 8u_{33} + 0.000 0u_{36} + 0.999 7u_{37}$$

由此可得相关性较弱(近似相互独立)的模型指标共 n 个。消除相关性后的数据记为 $Z = (z_{ij})_{m \times n}$, 显然 $z_{ij} > 0$ 。

3 模型构成

3.1 确定积分模型结构

为评估犯罪嫌疑人的危险程度,对犯罪嫌疑人进行分类分级,建立如下乘积形式的非线性积分模型:

$$S_i = \prod_{j=1}^n z_{ij}^{\beta_j}, i = 1, 2, \dots, m (\beta_j \geq 0), \quad (1)$$

式中: S_i 表示第 i 个嫌疑人的危险积分, β_j 是第 j 个指标的权重, β_j 待定。

3.2 参数拟合模型

为求得模型的最优参数,即使得模型特征指标重要性最优,以经由乘积积分模型(1)计算推送出的一级危险嫌疑人与近期抓获人员的比对人数为优化目标,确定参数拟合模型。

对 M 中的犯罪嫌疑人按(1)计算积分值 Y_i 并按积分值的高低排序,排序后的嫌疑人编号为 $((1), (2), \dots, (m))$, 显然 $S_{(1)} \geq S_{(2)} \geq \dots \geq S_{(m)}$; 给定一级危险嫌疑人比例 α (考虑警力情况及 M 的规模), 将积分排名位于前 α 的嫌疑人取出记作 P , 即 $|P|$ 约为 $[am]$ 且 $P \subseteq M$; 将近期抓获的在库嫌疑人所构成的集合记作 $Q, Q \subseteq M$; 记 $I(\beta_1, \beta_2, \dots, \beta_n) = |P \cap Q|$ 表示权重 $(\beta_1, \beta_2, \dots, \beta_n)$ 下近期抓获人员与模型(1)推送人员的吻合度,以此作为目标进行优化确定最优参数。即:

$$\begin{cases} \max_{\beta_j} I(\beta_1, \beta_2, \dots, \beta_n) \\ s. t. \quad S_i = \prod_{j=1}^n z_{ij}^{\beta_j}, i = 1, 2, \dots, m \\ \beta_j \geq 0, j = 1, 2, \dots, n, \end{cases} \quad (2)$$

求解所得 $\beta^* = (\beta_1^*, \beta_2^*, \dots, \beta_n^*)$ 为模型最优特征指标重要性。

参数拟合模型(2)的目标函数非解析,无法用求梯度信息的数值优化方法直接求解。模拟通过对过程或系统进行模仿为系统的指标提供一个统计估计,虽然结果并不精确,但其也许是唯一能够获得问题答案的方法^[14]。Monte Carlo 模拟可以不考虑目标函数的复杂性以及设计变量的维数,特别适用于目标函数复杂、变量多和严重非线性优化问题^[15]。在求解参数拟合模

型(2)的过程中, Monte Carlo 算法用于控制设计过程,即筛选模型参数,并设计合理的算法终止条件,根据判断结果对搜索区域做调整。图 1 给出了利用 Monte Carlo 算法和 MATLAB 进行优化的具体流程。

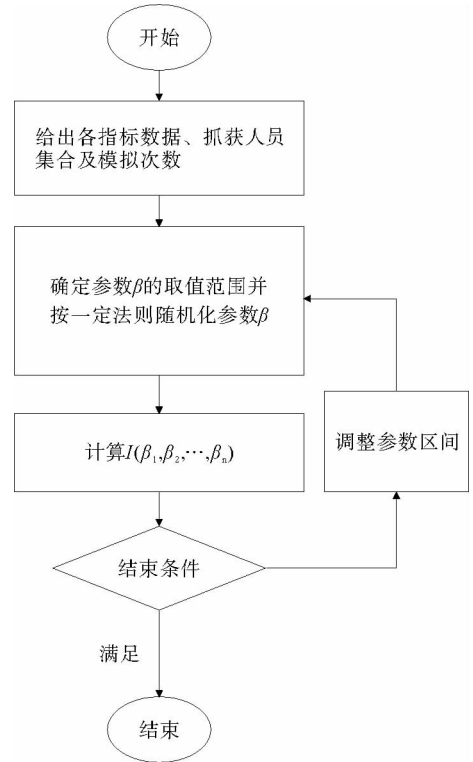


图 1 参数优化设计流程

Fig. 1 Parameters Optimization design process

3.3 近似最优参数的犯罪预警的乘积积分模型确定

通过对(3.2)参数拟合模型的求解,将求得的近似最优参数代入(3.1)即可确定近似最优参数的犯罪预警的乘积积分模型,即:

$$S_i = \prod_{j=1}^n z_{ij}^{\beta_j^*}, i = 1, 2, \dots, m. \quad (3)$$

4 模型求解

4.1 求解参数拟合模型的 Monte Carlo 算法

Monte Carlo 算法求解参数拟合模型(2)的基本思路是:以一定概率模型随机化参数变量,通过目标函数的抽样试算获得参数变量的统计特征,以样本均值作为参数变量的近似值^[16]。具体实现步骤为:

- (1) 产生 $[0, 1]$ 区间服从均匀分布的 $K \times n$ 阶随机数矩阵 $\gamma_{K \times n}$;
- (2) 每个指标权重 $\beta_j (j=1, 2, \dots, n)$ 取值范围为 $[t^{(l)}, T^{(l)}] (l=1, 2, \dots)$, 并给定初始值 $[t^{(1)}, T^{(1)}]$;
- (3) 以一定步长生成 K 组随机试验参数 $\beta_k^{(l)} = (\beta_{1k}^{(l)}, \beta_{2k}^{(l)}, \dots, \beta_{nk}^{(l)})$, $k=1, 2, \dots, K \quad \beta_{jk}^{(l)} = t^{(l)} + \gamma$

$(k, j) \cdot (T^{<l>} - t^{<l>}) (j=1, 2, \dots, n; k=1, 2, \dots, K)$;
 (4) 基于 $\beta_k^{<l>}$ 计算 $I(\beta_1, \beta_2, \dots, \beta_n) = |P \cap Q|$, 并排列对应的 $I_k^{<l>} = (\beta_{1k}^{<l>}, \beta_{2k}^{<l>}, \dots, \beta_{nk}^{<l>})$, $k=1, 2, \dots, K$, 记排名第一的目标值为 $I^{<l>}$, 并记录对应的 $\beta_k^{<l>} = (\beta_{1k}^{<l>}, \beta_{2k}^{<l>}, \dots, \beta_{nk}^{<l>})$;

$$(5) \text{ 当 } \sigma = \sqrt{\frac{1}{d} \sum_{r=l-d+1}^l \left(I^{<r>} - \frac{\sum_{r=l-d+1}^l I^{<r>}}{d} \right)^2} \leq \epsilon \text{ (} \epsilon \text{ 为}$$

人为给定的数)时, 过程结束, 返回 $\beta^* = \frac{\sum_{k=1}^l \beta_k^{<l>}}{d}$; 否

则返回第(2)步, 调整 $\beta_j (j=1, 2, \dots, n)$ 的取值范围: $\frac{T^{<l+1>}/t^{<l+1>}}{T^{<l>}/t^{<l>}} = \omega (\omega > 1)$ 且 $t^{<l+1>} = t^{<l>}$, 重新计算(3)。

4.2 实验结果及分析

将本文提出的方法用于某市的相关研究, 并与现有积分模型的预测结果对比, 模型相关参数如下: 给定 $K=1\ 000, \alpha=10\%$, 初始值 $t^{<l>} = \frac{1}{3}, T^{<l>} = 2, d=4$,

$\omega = \frac{4}{3}, \epsilon = 0.5$ 。模拟 N 次计算近似最优参数值, 并计

算模型的准确率, 记准确率 $\eta = \frac{\max I(\beta_1, \beta_2, \dots, \beta_n)}{|Q|}$, 现

有积分模型与本文构建的乘积积分模型准确率分别记为 η_0, η_1 , 模拟结果如表 3:

表 3 实验结果

Table 3 Experimental result

$ Q $	$ P $	η_0	η_1	有效提升 Effective promotion
1034	2013	24.2%	25.9%	1.7%

求得的近似最优参数为:

$\beta^* = (3.630\ 7, 0.993\ 9, 3.151\ 1, 0.854\ 1, 3.068\ 9, 3.548\ 6, 4.142\ 0, 3.446\ 8, 2.067\ 3, 0.629\ 6, 4.706\ 3, 0.984\ 8, 4.294\ 7, 3.016\ 4, 3.589\ 6, 0.577\ 1, 4.494\ 1, 0.377\ 3, 4.568\ 5, 0.551\ 4, 1.875\ 1, 0.615\ 9, 1.543\ 6, 2.089\ 8, 0.643\ 2, 3.831\ 4, 1.580\ 4, 1.281\ 3, 1.690\ 5, 0.635\ 2, 0.834\ 5, 1.833\ 6, 1.444\ 9)$;

将近似最优参数 β^* 代入模型(3.3)计算各嫌疑人的危险积分, 根据危险积分的高低对犯罪嫌疑人分类分级, 评估嫌疑人员的危险程度。

可见, 本文构建的乘积积分模型将嫌疑人的指标表现与特征指标的重要性分开来刻画, 根据不同的量化方式来确定嫌疑人指标表现的积分值, 利用 Monte Carlo 方法确定模型的近似最优参数, 即近似最优的特

征指标重要性, 解决了现有模型确定积分值的主观性; 且乘积积分模型对比目前投入使用的积分模型具有更好的准确率, 一旦模型指标确定, 算法就可按约束条件进行参数拟合求出模型的近似最优参数, 程序简单, 可操作性强。

5 结语

犯罪预警的乘积积分模型通过数据分析、数据处理及参数拟合, 对指标的重要性以及嫌疑人的指标表现分别进行刻画, 运用 Monte Carlo 方法求得模型的近似最优参数, 即模型最优特征指标重要性, 较现有积分模型的准确率有提升。并且积分模型对数据质量的要求较低, 给定数据信息, 只要确定模型的指标集, 就可以通过指标数量化、参数拟合计算嫌疑人的危险程度, 根据实际警力条件及犯罪嫌疑人数据规模确定高危嫌疑人集合, 更贴合实际。相关部门根据模型计算出的危险积分对犯罪嫌疑人分类分级进而做出相应的侦案决策, 在警力不足的情况下, 为提高侦案效率奠定了基础。

参考文献:

- [1] 高玉龙. 高危人群的识别和控管研究综述[J]. 吉林公安高等专科学校学报, 2010(2): 5-12.
Gao Yulong. Identify and control research of high-risk groups[J]. Journal of Jilin Public Security College, 2010(2): 5-12.
- [2] 邢轶清. 公安情报分析系统的设计与实现[D]. 内蒙古: 内蒙古大学, 2007.
Xing Yiqing. The design and implementation of the public security intelligence analysis system[D]. Mongolia: Inner Mongolia university, 2007.
- [3] 姜启源. 数学模型[M]. 北京: 高等教育出版社, 2003.
Jiang Qiyuan. Mathematical model[M]. Beijing: Higher Education Press, 2003.
- [4] 杨建军, 张勇坚. 温特斯法在 110 警情预测中的应用初探[J]. 云南警官学院学报, 2009(6): 45-50.
YANG Jianjun, ZHANG Yongjian. Application of winter methods to 110 policing alert prediction[J]. Journal of Yunnan Police Officer Academy, 2009(6): 45-50.
- [5] Mena J. Investigative Data Mining for Security and Criminal Detection[M]. UK: Butterworth-Heinemann, 2003.
- [6] Fei B K L, Eloff J H P, Olivier M S, et al. The use of self-organising maps for anomalous behaviour detection in a digital investigation[J]. Forensic Science International, 2006, 162(s1-3): 33-37.
- [7] Ruijuan Hu. Data mining in the application of criminal cases based on decision tree[J]. International Journal of Engineering Sciences, 2013, 2(2): 24-27.
- [8] Wang H, Wang J, Zheng T. Analysis of decision tree classification algorithm based on attribute reduction and application in criminal behavior[C]. Shanghai: Computer Research and Development (IC-CRD), 2011 3rd International Conference on IEEE, 2011: 27-30.
- [9] 汪明亮. 犯罪生成模式研究[M]. 北京: 北京大学出版社, 2007.

- Wang Mingliang. Crime study the model[M]. Beijing: Peking University Press, 2007.
- [10] 王岩. Monte Carlo 方法应用研究[J]. 云南大学学报: 自然科学版, 2006, 28(S1): 23-26.
- WANG Yan. Monte Carlo Method Applied Research[J]. Journal of YunNan University(Natural Science), 2006, 28(S1): 23-26.
- [11] 李亮, 吴瑞明. 消除评价指标相关性的权值计算方法[J]. 系统管理学报, 2009, 18(2): 221-225.
- LI Liang, WU Ruiming. A study on eliminating correlation of evaluatin index and obtaiing modified weights[J]. Journal of Systems & Management, 2009, 18(2): 221-225.
- [12] 徐祥发, 肖人彬. 评价指标相关性的消除方法研究[J]. 系统工程理论与实践, 2002, (11): 1-5.
- XU Xiangfa, XIAO Renbin. An approach of eliminating correlation of assessment-index[J]. System Engineering Theory and Practice, 2002, (11): 1-5.
- [13] 张大陆, 杨哲, 姚进. 分布式电子商务中服务评价指标相关性消除方法[J]. 同济大学学报(自然科学版), 2006, 34(3): 401-404.
- ZHANG Dalu, YANG Zhe, YAO Jin. Research on elimination the correlation of service evaluating index in distribution e-commerce[J]. Journal of TongJi University(Natural Science), 2006, 34(3): 401-404.
- [14] 刘东平. 随机物流模型的建模及求解方法研究[D]. 北京: 北京交通大学, 2005.
- Liu Dongping. Stochastic Modeling and Solving Method of Logistics Model Research[D]. Beijing: Beijing jiaotong university, 2005.
- [15] 陈立周. 工程离散变量优化设计方法—原理与应用[M]. 北京: 机械工业出版社, 1989.
- Chen Lizhou. Discrete Variable Optimization Design Method, Principle and Application[M]. Beijing: Beijing Mechanical Industry Press, 1989.
- [16] 黄平, 孟永钢. 最优化理论与方法[M]. 北京: 清华大学出版社, 2009.
- Huang Ping, Meng Yonggang. Optimization Theory and Method [M]. Beijing: Tsinghua University Press, 2009.

Research on Product Integral Model of Early Warning

WANG Qi, CAO Sheng-Shan, CHEN Pi-Wei, YAN Pan-Pan, LV Ke-Bo, GAO Xiang
(School of Mathematical Sciences, Ocean University of China, Qingdao 266100, China)

Abstract: This paper studies the classification of the criminal suspects by establishing mathematical model. With increasingly complex crime situation, the difficulty of public security and the social management and the cost has increased. The current domestic relevant departments mainly uses the integral model for the suspect controls. In view of the problems of the existing integral model of early warning which the functional departments concerned have been put into use, through the intelligence data of criminal suspects for data analysis and data processing, after extracting characteristic index and selecting the appropriate way to quantify indexes, the product integral model of early warning is constructed, and the solution of the parameters in the product integral model is formulated as a parameter fitting problem. As the objective function is non-analytic in parameter fitting problem, the approximate optimal parameters are determined by using Monte Carlo method. To determine the alarm level of criminal suspects by calculating the points of them, and then to focus on the classification of criminal suspects. Numerical experiments show that, compared with the existing integral model of early warning, the product integral model constructed in this paper has better coincidence, which lays a foundation for improving the efficiency of investigation.

Key words: crime early warning; product integral model; classification; parameter fitting; Monte Carlo method

责任编辑 陈呈超