文章编号:1674-7070(2010)06-0514-05

连环犯罪的时空预测方法

缪丽娟[1] 刘强[1] 王一舟[1] 陈昌春[1]

摘要

基于英国的连环犯罪调查数据,使用 ArcGIS9.2 软件对犯罪地点进行缓冲区分析,将全范围内的可能犯罪地点进行重分类,然后采用 Rossmo 公式建模预测分析犯罪地点,得到罪犯居住地的可能性分布,最后通过时间序列分析方法对下次行凶的可能时间和地点进行了预测.

关键词

缓冲区分析;反距离加权;时间序列分析

中图分类号 O29

文献标志码 A

收稿日期 2010-04-26

作者简介

缪丽娟,女.871909771@qq.com

1 南京信息工程大学 遥感学院,南京,210044

0 引言

Introduction

连环犯罪性质恶劣,对于社会安定造成的影响也较大.连环犯罪的多次犯罪之间或多或少地存在着一些联系,这为我们提供了进行分析预测的条件.犯罪地图是以空间地理信息为参照,操作与处理犯罪数据,以可视化形式显示、输出给特定用户有用信息的过程,是一种有效的情报分析工具,但当前在我国尚未建立起独立的学科[1],而用地理信息技术与数学方法相结合研究连环犯罪尤为薄弱[2].本文基于时空序列模型对连环犯罪的案例进行研究,为侦破连环杀人案件提供了 3 种预测方法,最后通过实例来进行分析说明.首先基于 GIS 地理轮廓构建,运用 ArcGIS9.2 软件对案发地点进行缓冲区分析,得到犯罪的地理轮廓,运用 Rossmo 公式计算罪犯各地居住概率,生成概率分布图,并与犯罪区域地图叠合分析,预测罪犯可能的居住区域.根据罪犯作案的时间间隔以及犯罪地点的统计数据,采用时空序列方法预测下一次犯罪的可能时间和地点,是对犯罪制图学的一次有益的探索与尝试.由于上述方法大都由过去的经验知识建立,存在一些不足之处,但通过叠合分析后可以减小可能误差,使结果更加精确.

1 资料来源与假设

Data resources and assumption

本文研究和讨论的案犯是连环案件,案犯是 1981 年被裁定为犯案 21 起、谋杀 13 名妇女和攻击数人罪的英国约克夏连环杀手彼得.通过 Wikipedia 获得了其犯罪的统计资料和大量的背景资料,并制作了统计图表用于直观展示.

为了方法的展开,本文的研究基于下列假设:

1)考虑时间地点,忽略其他(如犯罪心理,交通治安条件等)因素;

2)案件的性质为连环杀人案件;

3)假设罪犯在某一个较大的范围内重复作案,并且在短时间内不会移动到范围之外;

4)假设罪犯作案动机相对稳定,且惯于近距离作案.

2 连环犯罪地点预测方法与结果
Serial crime forecast method and result

如图1所示,为了便于分析罪犯犯案地点的规律,将West Yorkshire的地图按照城市边界的划分为5个区域,分别以A、B、C、D、E标注,区域A为Bradford,区域B为Leeds,区域C为Halifax,区域D为Huddersfield,区域E为Manchester. 根据罪犯的背景资料以及犯罪记录将犯罪的地点标注在上述的地图上,罪犯袭击的地方以a、b、c、d、e、f、g、h标注,发生命案的地点以1、2、3、4、5、6、7、8、9、10、11、12、13标注.

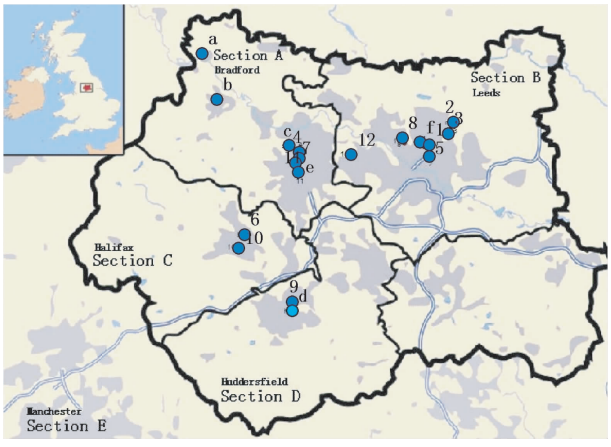


图1 犯罪地点分布
Fig. 1 Distribution map of crime places

2.1 犯罪区域地理轮廓的定性描述
利用ArcGIS9.2软件,对案发区域约克郡地图进行几何校正. 首先根据约克郡经纬度轮廓建立经纬网图层[3],然后对其进行投影定义,最后以地图4个角点作为控制点进行几何精校正. 将校正后图层与根据案发地点经纬度生成的点图层叠加,使用ArcGIS的缓冲区分析模块对点图层作缓冲区分析(图2),建立描述犯罪地点的地理轮廓.

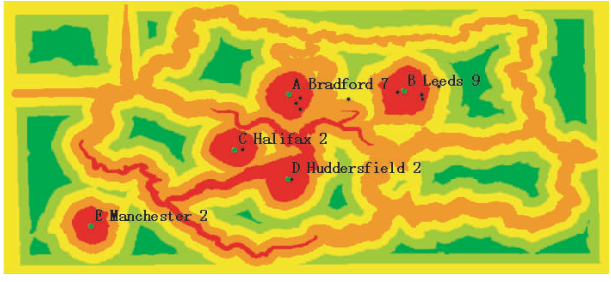


图2 缓冲区分析
Fig. 2 Buffer analysis

图2中,地理轮廓直观地描述了犯罪地点高发区,红色区域表示危险警戒区,应该加强警力,其余色块为相对安全的区域.

2.2 罪犯居住点的概率分布与计算
2.2.1 计算方法
将实际的连续地区离散成栅格单元以矩阵形式存储,计算每个小区域罪犯居住的概率值(矩阵中栅格单元的值). 根据假设说明采用反距离加权进行处理[4-5]:

$$p_{i,j} = k\sum_{n=1}^{c}\left[\underbrace{\frac{\phi}{\left(|X_i - x_n| + |Y_i - y_n|\right)^f}}_{\text{term1}} + \underbrace{\frac{(1-\phi)B^{g-f}}{\left(2B - |X_i - x_n| + |Y_i - y_n|\right)^g}}_{\text{term2}}\right]. \quad (1)$$

式(1)中:(X_i, Y_i)表示图像中每个小栅格的概率值;(x_n, y_n)表示第n个犯罪现象;term1部分表示罪犯在靠近作案地点藏身的几率,与距离的f次方成反比;term2部分表示以作案点为圆心,B距离为半径的圆区域,罪犯在圆内部居住的可能性很小,在圆区域外居住概率可通过反距离加权估计.

表1 参数定义说明
Table 1 Parameter definition

参数	定义
$P_{i,j}$	罪犯居住在编号为(i,j)的小区域a时候的概率
k	随机影响因素(除时间、地点外)的权重系数.
c	案件总数
ϕ	权重调节参数
f,g	与距离相关的指数,用以调节距离的影响程度.
(X_i, Y_i)	当整体连续地区以离散栅格存储于计算机中时,原地区小区域a在栅格矩阵中所处的行列号
(x_n, y_n)	第n次案件发生地的行列号
B	缓冲区大小,即在作案地点方圆B距离的范围内,罪犯居住的可能性很小.

2.2.2 概率分布计算
在Matlab环境下读入约克郡的地图,利用size()函数计算出图像矩阵大小为531×721. 将约克郡的地图栅格化,划分成为9×12的大块,每一块大约有60×60个栅格单元从而在地图上建立起一个参考坐标系,如图3所示. 分别统计各个栅格单元的犯罪案件数目.

对于犯案的块区中,取一个栅格单元,认为在该块区所犯的案件均在此栅格单元发生,用栅格单元来

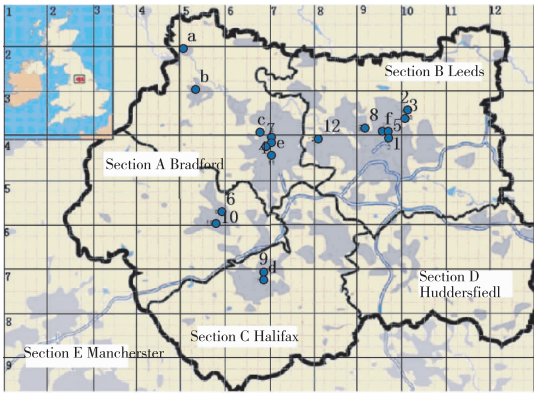


图 3　实际作案地点

Fig. 3　Actual crime places

代表整个块区,即所谓的“案发重心”,将此栅格单元的行列坐标值代入式(1)中作为(x_n, y_n)进行计算.

若单纯依照公式,当$(X_i, Y_j) = (x_n, y_n)$,即所计算地区为案发地区时,分母为 0,因此在计算时应稍作处理,可以设定 $p_{i,j} = 0$,认为罪犯肯定不会居住在此地. 对缓冲区距离,可预先设定一个值,而后根据实际验证加以改进,本文将初始值设为 10(即 10 个栅格单元).

Matlab 计算结果如图 4 所示.

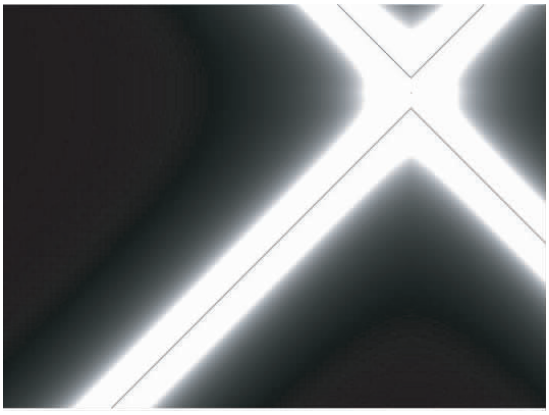


图 4　罪犯各地居住概率

Fig. 4　Living probability of the criminals

图 4 中,随着颜色亮化,各地区罪犯藏身概率逐渐增大. 图 4 所示的“十”字量化地带大致能表示出图 3 中案发地点分布特征,这与公式(1)中反距离加权项 term1 相应;在亮化中心处出现的“×黑线”表示在案件发生频率极高的地区,罪犯居住的概率较低,这与公式(1)中的缓冲区项 term2 相应,并且在案例中也予以证实——罪犯在 B 地作案最为频繁,可居住在邻近的 A 地区.

2.3　居住地的概率分布与实际区域叠合分析

将图 1 与案发区域地图进行叠置分析,如图 5 所示.

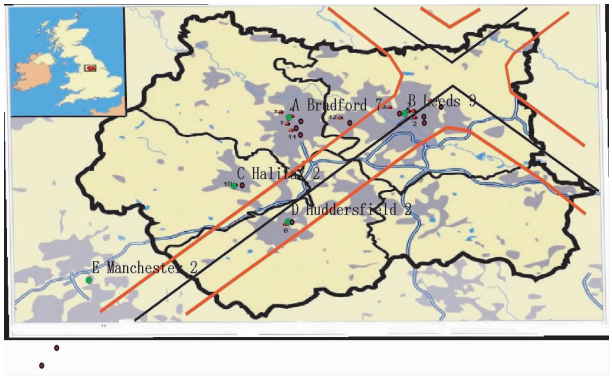


图 5　概率分布与实际区域叠合

Fig. 5　Congruence of living probability and actual area

图 5 显示罪犯作案区域(黑边为地图校正痕迹),图中紫色斑块为主要居民区,各点为罪犯实际作案地点,“十”字红线以内区域为 Rossmo 概率模型计算出的罪犯居住几率高值区,这和缓冲区模型得出的各地案发几率图是相一致的. 据此推想罪犯很有可能居住在 B 城镇(图中 Leeds),但根据罪犯居住在犯罪高发区的概率很小的假设(图 5 红线中间的两条黑线,表示案发高频区,罪犯居住几率很小),罪犯居住在 B 区的可能性减弱. 按照就近作案的假设,罪犯居住地案发几率较高并且距离作案高频区较近,罪犯极有可能居住在 A 区(Bradford),因为通过连接各主要案发地点,发现案发率较高,而距离各案发地点较近的只有 A 地区(表 2),并且 A 地居民区距离理论所得的居住概率高值区域距离也很近.

表 2　假想居住地与实际作案地实际距离

Table 2　Distance between actual crime places and assumed criminal living places

实际作案地	假想居住地				
	A	B	C	D	E
A	0	0.24	0.17	0.16	0.58
B	0.24	0	0.34	0.28	0.65
C	0.17	0.34	0	0.11	0.46
D	0.16	0.28	0.11	0	0.50
E	0.58	0.65	0.46	0.50	0
L_T	1.15	1.52	1.08	1.05	2.19

注:L_T 为假想居住地与各作案地距离之和

3 连环犯罪时间序列分析预测与结果
Time sequence analysis prediction and result of serial crimes

时间序列预测方法是用于预测时序发生事件的非常有效的方法,但此法在国内犯罪预测方面还用得很少. 关于时间间隔序列的预测分析,首先使用Excel 描绘出时间间隔以及罪犯的犯罪次数之间的关系图,然后为关系图添加趋势线确定应该使用哪种曲线可以更好地对这个时间间隔序列进行描述. 通过对曲线形状以及走向初步确定可以使用高次指数以及移动平均法,经过尝试发现移动平均法可以更好地对上述曲线进行拟合. 时间间隔统计结果见表3,拟合效果见图6.

表3 罪犯作案时间间隔统计结果

Table 3 Statistical result of crime time intervals

犯罪次序	日期	时间间隔/d	修正后间隔/d	权重
1	1975-07-05	41	41	0.030 438
2	1975-08-15	12	12	0.008 909
3	1975-08-27	64	64	0.047 513
4	1975-10-30	82	82	0.060 876
5	1976-01-20	110	110	0.081 663
6	1976-05-09	272	95	0.070 527
7	1977-02-05	77	77	0.057 164
8	1977-04-23	64	64	0.047 513
9	1977-06-26	14	14	0.010 393
10	1977-07-10	83	83	0.061 618
11	1977-10-01	74	74	0.054 937
12	1977-12-14	38	38	0.028 211
13	1978-01-21	10	10	0.007 424
14	1978-01-31	105	105	0.077 951
15	1978-05-16	323	95	0.070 527
16	1979-04-04	151	151	0.112 101
17	1979-09-02	351	95	0.070 527
18	1980-08-18	37	37	0.027 468
19	1980-09-24	42	42	0.031 180
20	1980-11-05	12	12	0.008 909
21	1980-11-17	46	46	0.034 150

注:犯罪资料基于 Wikipedia 的资料整理[5-6].

本文所引用的案例为约克夏连环杀手彼得,他被裁定犯案21起,其中包括13起谋杀和8起袭击事件. 根据罪犯的犯罪时间和后一次犯罪时间的差值来确定犯罪时间间隔,彼得于1981年1月2日犯罪时被抓获,所以第21次犯罪的时间间隔为其与1980年11月17日之差.

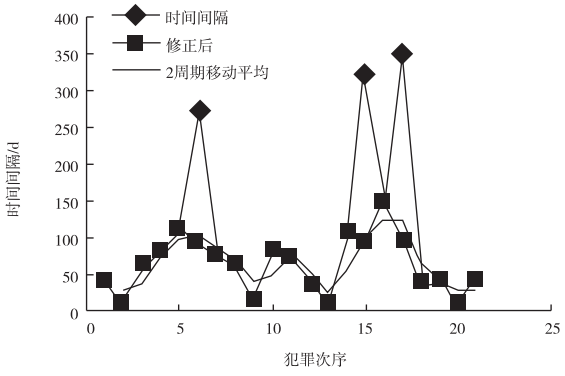


图6 犯罪时间间隔的拟合效果

Fig. 6 Fitted result of crime time intervals

为了更好地对曲线进行拟合,选择带权的移动平均法进行拟合

$$F_t = W_1 \cdot A_1 + W_2 \cdot A_2 + \cdots + W_n \cdot A_n,$$
$$n = 1,2,3,\cdots,t-1.$$

式(2)中:F_t 为预测的 t 次犯罪的时候距离上一次犯罪的时间间隔;A_n 为第 n 次犯罪以后距离下一次犯罪的时间间隔. w_n 第 n 项的权重,本模型中通过每次犯罪和下次犯罪之间的时间间隔与罪犯之前所有犯罪的时间间隔的比值来决定,满足

$$w_i = \frac{A_i}{\sum_{i=1}^{n} A_i}, \quad w_1 + w_2 + \cdots + w_n = 1.$$

图像上存在的一些时间间隔较大的特异的点,通过相关资料的研究发现在特异的点附近罪犯存在一些随机因素,比如搬到一个新的城市,罪犯需要时间去熟悉当地的环境,罪犯得到了一份比较稳定的工作造成罪犯的心理动机得到了缓解,罪犯的亲人的去世等,所以需要对特异的点的数据采用平滑处理,本案例中采用取平均值的方法进行处理.

4 结论和讨论
Conclusion and discussion

1) 本文的方法可以定性地描述连环犯罪罪犯的地理轮廓图,对下次犯罪发生的时间和地点做出大致的预测,在案件高发地区设置红色警戒区.

2) 由于连环犯罪的时间、地点存在一些随机影响因素,也可能造成预测偏差,但是可以得到一个大致的时间间隔,可以指导警方组织警力进行抓捕.

3) 本文所引案例的数据量不足,但是在详细案

情数据的支持下,其预测的精度可望进一步提升.

由于我国在犯罪制图学方面的起步较晚,用地理信息技术与数学方法相结合研究连环犯罪尤为薄弱,因此,本文的研究不失为一次有益的尝试.

致谢:感谢夏大峰教授给予本文理论上的指导和方法上的建议!

参考文献
References

[1] 孙峰华,毛爱华. 犯罪地理学的理论研究[J]. 人文地理, 2003,18(5):70-74
SUN Fenghua, MAO Aihua. Study on the theories of criminal geography[J]. Human Geography, 2003, 18(5):70-74

[2] 贺日兴. 犯罪制图:概念、现状与趋势[C]//全国地图学与 GIS 学术会议论文集. 福州,2004
HE Rixing. Crime mapping: Concepts, status and trends [C]// National Cartography and GIS Conference. Fuzhou, 2004

[3] Wikipedia. Yorkshire ripper [EB/OL]. [2010-05-27]. http://en.wikipedia.org/w/index.php? title = Peter_Sutcliffe & action = edit & section = 3

[4] Rossmo K. Geographic profiling [M]. Boca Raton: CRC Press, 2000

[5] Wikipedia. Rossmo's formula [EB/OL]. [2010-04-16]. http://en.wikipedia.org/wiki/Rossmo%27s_formula

[6] Wikipedia. Peter sutcliffe [EB/OL]. [2010-05-27]. http://en.wikipedia.org/wiki/Yorkshire_Ripper

Space-time prediction of serial crimes

MIAO Lijun[1] LIU Qiang[1] WANG Yizhou[1] CHEN Changchun[1]
1 School of Remote Sensing, Nanjing University of Infornation Science & Technology, Nanjing 210044

Abstract With the increasing complexity of modern society, the occurrence of serial crimes climbed rapidly, which increased the detection difficulties for the police. Based on the survey data of serial crimes happened in Yorkshire UK, this paper analysed the buffer areas of crime places using software of ArcGIS 9.2, and established a model with Rossmo fomula to reclassify crime places. This research also gives the probable distribution of criminals' residence, and predicts the probable time and place of next crime by using the time sequence.

Key words buffer area analysis; inverse distance weight; time series analysis