



# Metodi statistici per la predizione della criminalità

Rassegna della letteratura su predictive policing  
e moduli di data mining

Andrea Di Nicola  
Giuseppe Espa  
Serena Bressan  
Maria Michela Dickson  
Andrea Nicolamarino





eCrime Working Papers

n. 02

*Metodi statistici per la predizione della criminalità*

*Rassegna della letteratura su predictive policing  
e moduli di data mining*

Andrea Di Nicola (Coordinatore scientifico)

Giuseppe Espa (Responsabile Analisi Statistiche Spaziali Predittive)

Serena Bressan (Project Manager)

Maria Michela Dickson (Analisi Statistiche Spaziali Predittive)

Andrea Nicolamarino (Junior Researcher)

ISSN 2284-399X

ISBN 978-88-8443-587-3

eCrime - ICT, Law & Criminology  
Facoltà di Giurisprudenza  
Università degli Studi di Trento  
Via G. Verdi, 53  
38122 - Trento  
0461 282336  
[www.ecrime.unitn.it](http://www.ecrime.unitn.it)

Le opinioni espresse nel presente rapporto di ricerca sono di responsabilità esclusiva degli autori e non riflettono necessariamente la posizione ufficiale dell'Unione europea.

Copyright immagine Cap. n. 2: "Mark Scott Spatny/Shutterstock"

Trento, Dicembre 2014

© 2014 eCrime - Università degli Studi di Trento



# Indice

## 00

Introduzione	1
Obiettivo e organizzazione del lavoro	1
Il progetto eSecurity e la sicurezza urbana predittiva	1

## 02

Le esperienze di polizia predittiva a livello internazionale	13
Il caso di Memphis in Tennessee e il software Blue CRUSH (USA)	13
Il caso di Los Angeles in California e il software PredPol (USA)	14
Il caso di Trafford - Greater Manchester (UK)	16
Concetti fondamentali	17

## B

Bibliografia	33
Capitolo 1	33
Capitolo 2	34
Capitolo 3	34

## 01

Cos'è il predictive policing?	5
Perché il crimine è prevedibile	7
Una strategia innovativa per il contrasto della criminalità	8
La lotta al crimine nella società dell'informazione	9
Concetti fondamentali	10

## 03

Metodi statistici per la previsione della concentrazione della criminalità	19
Hot spot analysis	20
Data mining	23
Metodi di regressione	24
Near repeat methods	26
Analisi spazio-temporale	26
Risk terrain analysis	29
Agent based modelling	30
Conclusioni	31

# Introduzione



Questo working paper si inserisce nell'ambito dell'attività 5.1 del Workpackage n. 5 “Moduli di data mining per l'analisi criminologica” del progetto europeo “eSecurity - ICT for knowledge-based and predictive urban security” (nel seguito anche eSecurity), co-finanziato dalla Commissione europea (Programma ISEC 2011 “Prevention of and Fight against Crime” della DG Home Affairs - HOME/2011/ISEC/AG), primo progetto al mondo di sicurezza urbana predittiva. Il progetto è coordinato dal gruppo di ricerca eCrime della Facoltà di Giurisprudenza dell'Università degli Studi di Trento, in partnership con la Questura di Trento, la Fondazione Bruno Kessler e il Comune di Trento.

## Obiettivo e organizzazione del lavoro

Il fine del presente lavoro è di elaborare una guida essenziale per chiunque desideri conoscere e comprendere la natura ed il funzionamento dei più recenti progetti di *predictive policing* (polizia predittiva) e di *predictive urban security* (sicurezza urbana predittiva), come eSecurity. I destinatari di questo *working paper* sono principalmente ricercatori, agenti delle forze dell'ordine e funzionari degli enti pubblici interessati al tema in oggetto. Per raggiungere questo obiettivo, dopo aver definito il concetto di polizia predittiva nei suoi aspetti basilari (capitolo 1) e aver descritto le principali esperienze internazionali nelle quali si sono articolate tali attività (capitolo 2), saranno delineati i moduli predittivi statistici più utilizzati a livello mondiale, nell'ambito della prevenzione e del contrasto della criminalità e della gestione della sicurezza urbana (capitolo 3)<sup>1</sup>.

<sup>1</sup> La parte della ricerca di eSecurity a cui questo documento si riferisce è stata portata avanti da Andrea Nicolamarino (junior researcher ad eCrime), Maria Michela Dickson (ricercatrice ad eCrime) e Serena Bressan (project manager di eSecurity e ricercatrice ad eCrime), sotto la supervisione di Andrea Di Nicola (coordinatore scientifico di eSecurity e professore aggregato di criminologia presso la Facoltà di Giurisprudenza dell'Università degli Studi di Trento) e di Giuseppe Espa (responsabile dell'area “Analisi statistiche spaziali predittive” di eSecurity e professore ordinario di statistica economica presso il Dipartimento di Economia e Management dell'Università degli Studi di Trento).

## Il progetto eSecurity e la sicurezza urbana predittiva

Il progetto eSecurity si basa sui principi delle teorie razionali del crimine e della criminologia ambientale, secondo i quali la criminalità a livello urbano si concentra in alcuni “luoghi” (punti, strade, zone) e la vittimizzazione passata è predittore di quella futura. Questa concentrazione spazio-temporale di criminalità è dovuta ad una concentrazione spazio-temporale di opportunità, di cause, e queste vanno investigate per incidere sulla criminalità nelle città (Brantingham e Brantingham, 1991). Pertanto, la conoscenza dei punti “caldi” del territorio (“hot spot”) e delle opportunità criminali esistenti permette di poter valutare al meglio quali siano i fattori criminogeni da tenere in considerazione nell'elaborazione di politiche ed interventi di prevenzione e contrasto efficaci ed efficienti (Clarke, 1997; Wartell e Gallagher, 2012).

Seguendo questi presupposti teorici, eSecurity prende le mosse dalle esperienze pilota di *predictive policing*. Si tratta, in particolare, di progetti americani ed inglesi: l'esperienza dell'IBM insieme all'Università e alla Polizia di Memphis (USA), quella dell'Università della California Los Angeles e dell'Università della California Irvine con la Polizia di Los Angeles (USA) e, infine, il progetto del Jill Dando Institute of Security and Crime Science (University College of London) e della Polizia di Trafford, Greater Manchester (UK), che saranno presentate in seguito nelle pagine di questo documento. Per *predictive policing* si intende nel dettaglio l'attività di analisi dei dati di polizia sui crimini avvenuti in passato, della loro collocazione spazio-temporale (reati denunciati georiferiti) e delle ricorrenze riscontrate negli schemi di comportamento dei criminali, per prevedere i luoghi di futura concentrazione della criminalità sul ter-

**Il progetto eSecurity si basa sui principi delle teorie razionali del crimine e della criminologia ambientale, secondo i quali la criminalità a livello urbano si concentra in alcuni “luoghi”.**

## I concetti di disordine, vittimizzazione e insicurezza percepita sono, quindi, utili strumenti di diagnosi per comprendere i processi di sicurezza oggettiva e soggettiva in città.

ritorio, con il fine ultimo di allocare le risorse di polizia in modo ottimale (RAND, 2013).

Il progetto eSecurity, rispetto a queste esperienze straniere di polizia predittiva, effettua diversi passi avanti, al fine di sperimentare il nuovo modello di sicurezza urbana predittiva. Secondo questo nuovo modello teorico, le informazioni sulla vittimizzazione, sul disordine urbano ed altre variabili ambientali (ad esempio, illuminazione, traffico, clima) georiferiti, se letti in combinazione con i dati di polizia, possono evidenziare regole predittive in materia di sicurezza oggettiva e soggettiva, a supporto dell'azione di forze dell'ordine ed amministratori locali nella città. Questo nuovo paradigma per la prevenzione ed il contrasto del crimine ha ricombinato, in maniera del tutto innovativa grazie all'utilizzo delle tecnologie ICT, le conoscenze che provenivano dagli studi delle teorie razionali e della criminologia ambientale e la grande quantità di dati di cui possiamo disporre al giorno d'oggi nella società dell'informazione (Di Nicola e Bressan, 2014).

Il progetto europeo eSecurity ha l'obiettivo di sviluppare uno strumento ICT innovativo e georiferito (prototipo) per la raccolta dati, con lo scopo di migliorare le attività di gestione della sicurezza urbana e della prevenzione della criminalità e dei fenomeni di devianza in città. Il suo fine ultimo è assistere i decisori politici e le forze di polizia. In particolare, nell'area pilota del comune di Trento si stanno realizzando:

1. un database georiferito (eSecDB), concepito per immagazzinare dati su eventi criminali e informazioni su vittimizzazione, percezione della sicurezza, disordine urbano e altre variabili rilevanti (ad esempio, variabili socio-demografiche, informazioni su condizioni climatiche, traffico, trasporti pubblici, inquinamento);
2. un sistema informativo geografico (eSecGIS), che utilizza come input i dati provenienti da eSecDB, con capacità avanzate di generazione automatica di report, di visualizzazione di mappe di rischio e di sicurezza urbana predittiva;
3. un portale web (eSecWEB), per rafforzare la comunicazione e la collaborazione tra cittadini, amministrazioni locali e forze dell'ordine su politiche, iniziative e consigli su possibili comportamenti preventivi.

Per la predizione e la prevenzione delle future concentrazioni di criminalità e devianza, il sistema informativo geografico eSecGIS (prototipo), con i relativi algoritmi predittivi in esso inseriti, non si serve solo dei dati sui luoghi e sulle tempistiche degli eventi criminali passati georiferiti e anonimizzati, provenienti dallo SDI (Sistema di Indagine) del Ministero dell'Interno immagazzinati in eSecDB, ma utilizza altre variabili socio-demografiche e ambientali georiferite, derivanti anche dalla "Smart City", la città intelligente: ad esempio, il livello dell'illuminazione nei diversi quartieri, la situazione climatica, il traffico o i numeri relativi all'utilizzo dei trasporti pubblici (Di Nicola e Bressan, 2014).

Considerare queste variabili è fondamentale nella gestione della sicurezza urbana, perché nelle città esistono luoghi che sono considerati generatori o attrattori di criminalità di per sé (*crime generators/attractors/enablers areas*): ad esempio, esercizi commerciali, industrie, bar, banche, parchi o parcheggi. In queste aree, tendono a concentrarsi criminalità, disordine urbano e insicurezza. Il prototipo eSecGIS potrà aiutare a comprendere che intervenire su tali luoghi è cruciale per prevenire e ridurre i reati, diminuendo altresì lo spreco di risorse pubbliche. In tal modo, infatti, possono anche essere individuate strategie preventive ad hoc, mirate alla modifica dei comportamenti e alla pianificazione urbana (Brantingham e Brantingham, 1995; Lab, 2010).

Inoltre, questo strumento ICT tiene anche conto della concentrazione del disordine (o degrado) urbano fisico e sociale (reale e percepito), della vittimizzazione e dell'insicurezza percepita, quali predittori della concentrazione della criminalità e della devianza nel tessuto urbano (Nobili, 2003; Regione Piemonte, 2012). La paura della criminalità delle persone sembra crescere quando alla percezione del rischio di vittimizzazione si accompagna il disordine urbano. In questa categoria, si distingue tra fenomeni di disordine fisico (ad esempio, graffiti sui muri, rifiuti abbandonati, edifici in cattive condizioni) e sociale (ad esempio, presenza di tossicodipendenti, prostitute, vagabondi). Il senso di insicurezza di chi abita in uno spazio urbano tende ad intensificarsi con l'aumentare di questi segnali di degrado ambientale e, nello specifico, in corrispondenza di una più accentuata violazione delle regole riguardanti l'uso degli ambienti cittadini. I concetti di disordine, vittimizzazione e insicurezza percepita sono, quindi,

utili strumenti di diagnosi per comprendere i processi di sicurezza oggettiva e soggettiva in città (Wilson e Kelling, 1982; Chiesi, 2003).

Per raccogliere questi dati, nell'ambito del progetto eSecurity, saranno innanzitutto svolte quattro indagini di vittimizzazione a cadenza semestrale, denominate Indagini sulla sicurezza oggettiva e soggettiva nel comune di Trento: dopo il primo *round* d'indagine svoltosi ad ottobre 2013, il relativo questionario è stato somministrato ai cittadini ad aprile 2014 ed ottobre 2014 e lo sarà nuovamente ad aprile 2015. Lo scopo delle indagini è di raccogliere informazioni sui reati subiti negli ultimi sei mesi e nell'ultimo anno dai cittadini del capoluogo trentino, nonché sul senso di insicurezza e sui livelli di disordine urbano percepiti nel territorio comunale. Invece, la raccolta dei dati georiferiti sul disordine urbano fisico e sociale presente sul territorio di Trento sarà svolta tre volte ogni sei mesi (aprile 2014, ottobre 2014, marzo 2015) su base circoscrizionale, grazie alla Questura di Trento che rileverà le situazioni di degrado attraverso un device dedicato. I rilevatori utilizzeranno un'applicazione per la geo-localizzazione dei fenomeni di disordine sviluppata *ad hoc*.

In questo modo, attraverso il continuo dialogo tra il database eSecDB e il sistema informativo geografico eSecGIS, il progetto eSecurity cerca non solo di prevenire e prevedere le concentrazioni spazio-temporali della criminalità e della devianza a livello urbano, con un maggiore grado di precisione rispetto ai risultati raggiunti con le esperienze anglosassoni di polizia predittiva, ma anche di capire le cause dei reati, degli atti devianti, del disordine urbano e delle manifestazioni di insicurezza della popolazione (Di Nicola e Bressan, 2014). Dopo aver brevemente analizzato questo modello innovativo per la gestione della sicurezza a livello urbano, sviluppato nell'ambito di eSecurity, è doveroso fare un passo indietro, andando ad analizzare le esperienze di polizia predittiva esistenti a livello mondiale ed i metodi statistici che ne stanno alla base, attraverso un'analisi della letteratura di riferimento.

**La sicurezza urbana predittiva ha lo scopo di capire le cause dei reati, degli atti devianti, del disordine urbano e delle manifestazioni di insicurezza della popolazione.**

# 01

## Cos'è il predictive policing?



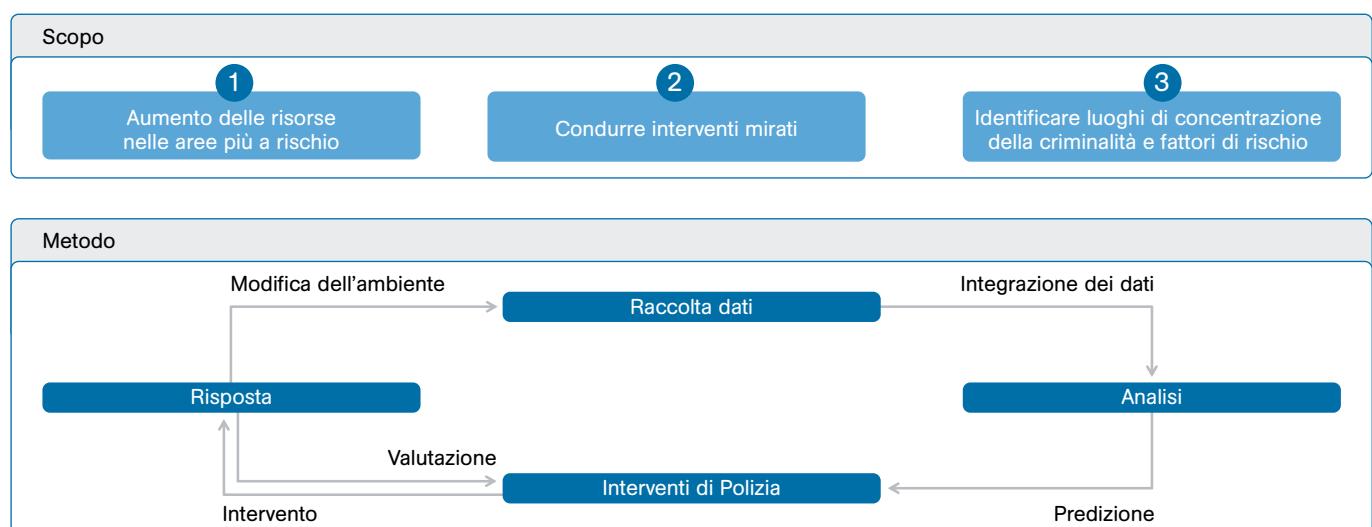
Il *predictive policing* (polizia predittiva) è una strategia di prevenzione del crimine e/o tattica di polizia che utilizza informazioni e sviluppa analisi avanzate per la previsione delle zone a più alta densità criminale in ambito cittadino, al fine di prevenire la criminalità futura (Uchida, 2009). Questa tipologia di prevenzione degli atti criminali attinge a piene mani alla criminologia ambientale, ovvero l'evoluzione ultima delle teorie criminologiche razionali, che studia come i target criminali si muovano nello spazio e nel tempo, riservando particolare attenzione alla distribuzione geografica del crimine e al ritmo delle attività giornaliere (Vezzadini, 2006). Tale teoria sostiene che gli eventi criminali hanno luogo al ricorrere di determinati fattori spazio-temporali, ovvero alla convergenza di delinquenti, vittime o obiettivi in contesti specifici, in un tempo e in uno spazio definiti (Brantingham & Brantingham, 1991).

Da questo *background* teorico, che postula, dunque, l'esistenza di schemi ricorrenti nella commissione di determinati atti criminali, deriva la possibilità, inedita per le forze dell'ordine, di interrompere il meccanismo causale del crimine, anticipandolo (Roufa, n.d.): se un modello esiste, infatti, è possibile sapere tempestivamente quando, dove e come saranno commessi i

crimini (Figura 1). Ciò, di conseguenza, incide sul funzionamento concreto della polizia predittiva: anni di dati e informazioni sugli eventi criminali vengono raccolti e catalogati in banche dati integrate e successivamente inseriti in un software che li analizza, col fine ultimo di trasformare tali informazioni dapprima in conoscenza sul dove e quando sia più probabile che avvenga un crimine e, conseguentemente, in una guida per la prevenzione (Beck & McCue, 2009).

Infatti, attraverso l'analisi dei dati relativi ai luoghi di maggiore concentrazione dei crimini avvenuti in passato (*hot spot*) individuati dal software e l'utilizzo di modelli predittivi (modelli statistici, moduli di *data mining*), le forze dell'ordine possono organizzare in maniera più efficiente le risorse a propria disposizione, posizionandosi meglio sul territorio cittadino. In questo modo, possono essere presenti nelle zone in cui è previsto che si verifichi un reato quel giorno e in quella fascia oraria. Di conseguenza, hanno la possibilità di condurre interventi mirati e specifici. Il tipo di operazione che la polizia porrà in essere potrà variare a seconda della finalità cui mira, se la semplice deterrenza o, piuttosto, la soluzione definitiva di una questione criminale. L'operazione può condurre a dei risultati che

**Figura 1.** Il metodo del *predictive policing*



Fonte: elaborazione eCrime su RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013, p. XVIII

**Il predictive policing è una strategia di prevenzione del crimine che utilizza informazioni e sviluppa analisi avanzate per la previsione delle zone a più alta densità criminale in ambito cittadino.**

presumibilmente ridurranno la criminalità in una determinata zona. Questi risultati vengono, quindi, inviati al database iniziale per essere ricompresi tra i dati e le informazioni da analizzare. E il ciclo si ripete (RAND, 2013). Allo stato attuale, le strategie di polizia predittiva possono essere classificate all'interno di quattro ampie categorie:

1. *metodi per la previsione della criminalità:* sono approcci che mirano a prevedere il luogo e il tempo in cui è più probabile che avvenga un crimine. Generalmente questi metodi predittivi comportano lo studio dei dati relativi ai crimini passati, in quanto l'idea sottintesa è che il passato sia il prologo di quel che accadrà. Chiari esempi di tale approccio sono i progetti statunitensi e britannici di *predictive policing*, tra le esperienze internazionali che indubbiamente hanno fornito i migliori risultati in ambito di gestione e contrasto del fenomeno criminale in ottica preventiva e predittiva;
2. *metodi per la previsione del futuro aggressore:* sono approcci che puntano ad identificare le classi di individui maggiormente a rischio di commettere un reato in futuro. Anche in questi casi, il database criminologico alla base delle analisi predittive è costituito essenzialmente dai record spazio-temporali dei crimini passati, cui, però, sono abbinate le informazioni relative ai criminali che li hanno commessi. Numerosi esempi di tale approccio si hanno negli Stati Uniti: in questi progetti, i ricercatori, dapprima identificano nel *background* del criminale una serie di fattori socio-demografici potenzialmente predittori di violenza; successivamente, elaborano un indice di rischio che viene utilizzato dalle forze dell'ordine per intervenire e/o prevenire i possibili casi di recidiva;
3. *metodi per l'elaborazione di un identikit criminale:* sono approcci che puntano a creare profili per ricercare gli autori di reati passati. Essi si basano su modelli in grado di elaborare le schede dei potenziali sospettati a partire dalle informazioni

ni conservate nei database più eterogenei: un esempio, in tal senso, è il britannico PNC (*Police National Computer*), un vasto database costruito per immagazzinare i dati relativi alle automobili in circolazione, che ha aiutato la polizia britannica nell'attività di prevenzione e repressione del crimine, anche grazie all'integrazione con il *National Automated Fingerprint Identification System*, un database contenente immagini delle impronte digitali di coloro che sono stati condannati o arrestati, e alla combinazione con l'*Automated Number-Plate Recognition*, un software che utilizza un computer a rete neurale per riconoscere i numeri delle targhe riprese da telecamere aeree (Lyon, 2003);

4. *metodi per la previsione della futura vittima:* similmente agli approcci precedenti, la strategia di polizia predittiva così sviluppata mira a identificare i gruppi o, in alcuni casi, gli individui più a rischio di vittimizzazione. Base di partenza è l'integrazione dei dati provenienti da diverse fonti, come, ad esempio, le informazioni delle indagini di vittimizzazione con il database in possesso delle forze dell'ordine; i risultati di tale operazione consentono alle forze di polizia di aumentare la vigilanza nelle zone frequentate dai soggetti più a rischio.

Nelle prossime sezioni, l'attenzione sarà focalizzata sui metodi rientranti nella prima delle categorie presentate, ovvero sulle esperienze internazionali che hanno conseguito i risultati più soddisfacenti nel contrasto della criminalità in un'ottica preventiva; in particolare, si farà riferimento ai progetti statunitensi e britannici di *predictive policing*.

**Anni di record su eventi criminali vengono raccolti in banche dati integrate ed analizzati da un software, col fine ultimo di trasformarli in guida per la prevenzione.**

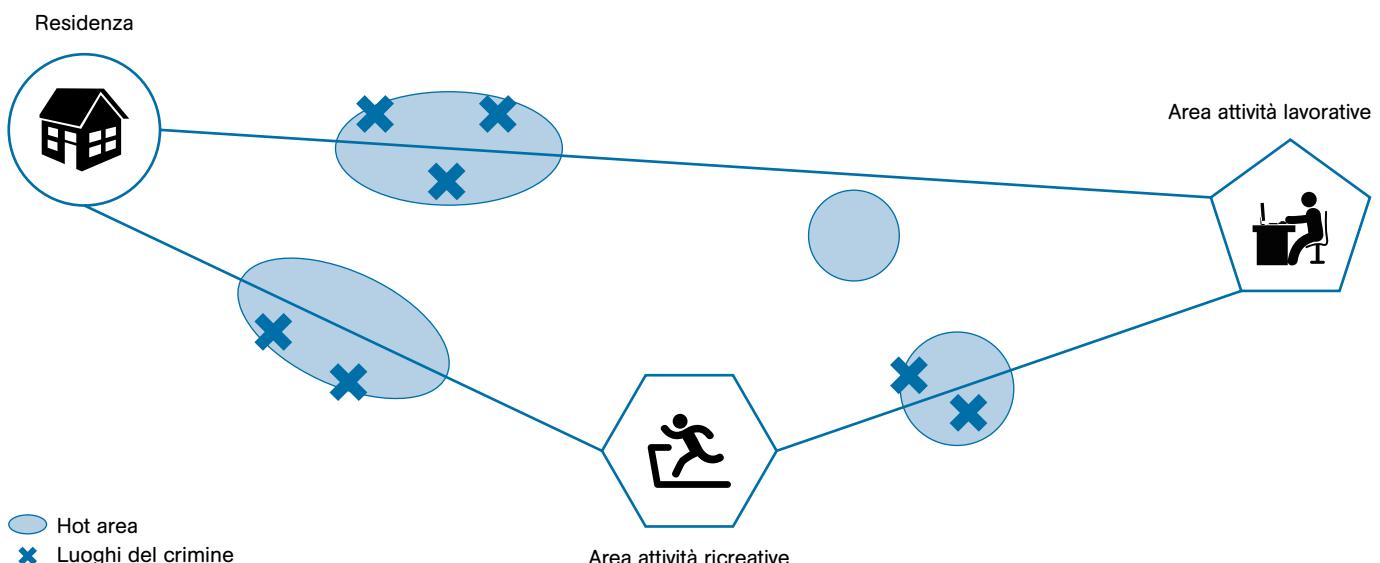
## Perché il crimine è prevedibile

Il fondamento teorico della polizia predittiva è rappresentato dalle teorie razionali e dalla criminologia ambientale, le quali ritengono che il crimine sia prevedibile sulla base dell'assunto che un criminale deciderà di commettere un delitto ogni volta il crimine sia altamente desiderabile e vi sia l'opportunità di commetterlo. Questa opportunità è data dal ricorrere di una serie di fattori, così come definiti dalle teorie razionali del crimine ed, in particolare, dalla "Teoria delle Attività di Routine" (Figura 2): la presenza di un autore motivato (*motivated offender*) e di "obiettivi/bersagli interessanti" (*suitable targets*), con la concomitante assenza del cosiddetto "guardiano capace" (*capable guardian*), cioè di una persona o comunque di un sistema (ad esempio, telecamere a circuito chiuso) che sia in grado di impedire che il crimine venga portato a compimento (Williams & McShane, 2002).

In particolare, le teorie razionali sottolineano alcuni aspetti:

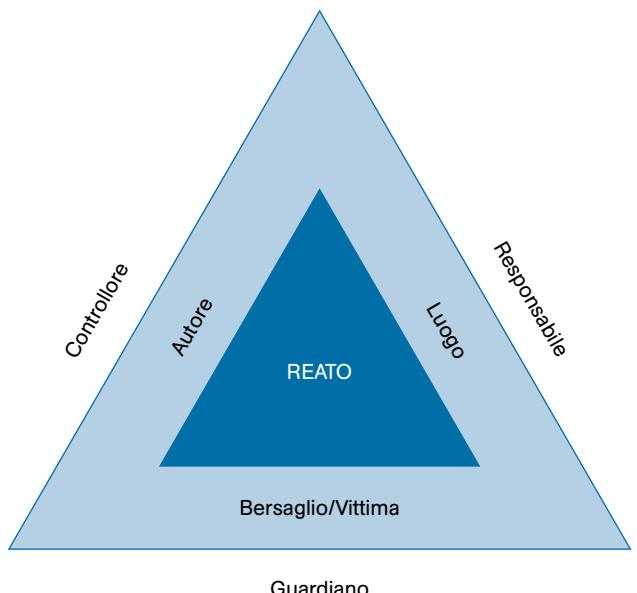
- quando le attività abituali quotidiane di criminali e vittime si incrociano, la probabilità che avvenga un crimine cresce esponenzialmente (Figura 3). Alcuni luoghi (ad esempio, bar, parchi, centri commerciali), infatti, sono generatori o attrattori di criminalità di per sé (i cosiddetti *crime generators/attractors*),

Figura 3. Lo schema di comportamento dei criminali



Fonte: elaborazione eCrime su K. Rossmo, *Geographic profiling*, Boca Raton, 2000

Figura 2. Il triangolo della criminalità secondo la "Teoria delle Attività di Routine"



Fonte: elaborazione eCrime su R. Clarke – J. Eck, *Problem solving e analisi criminale* (trad. it.), Trento, 2008, p. 27

**Le caratteristiche spazio-temporali influenzano il dove e il quando sia più probabile che avvenga un crimine.**

in quanto consentono la contemporanea presenza di criminali e vittime/obiettivi (Felson & Clarke, 1998);

- le caratteristiche spazio-temporali influenzano il dove e il quando sia più probabile che avvenga un crimine. Infatti, la criminalità non è diffusa uniformemente, ma si concentra in alcuni luoghi (i cosiddetti *hot spot*) e durante alcune ore della giornata o alcuni periodi della settimana/mese/anno (Sherman, 1995; Harries, 1999);
- i criminali prendono decisioni razionali sull'opportunità di perpetrare un reato considerando fattori come il luogo, la facilità di colpire il bersaglio e il rischio di essere scoperti (Felson & Cohen, 1979; Cornish & Clarke, 1986).

Da questi assunti deriva per le forze dell'ordine che rendono operativa la strategia del *predictive policing* la possibilità di orientare, in maniera più specifica e mirata, la propria azione preventiva, volta all'interruzione del meccanismo causale del crimine, mediante il rafforzamento della presenza degli agenti nei luoghi a più alta densità criminale. In questo modo, si vuole incentivare la funzione deterrente e, di conseguenza, rendere più rischiosa e meno vantaggiosa la commissione dell'atto criminale (Felson & Cohen, 1979; Cornish & Clarke, 1986).

**Il crimine è influenzato dalla routine quotidiana di vittime e criminali.  
Ad esempio viaggiare in metropolitana durante le ore di punta aumenta la possibilità di subire un furto di oggetti personali.**

## Una strategia innovativa per il contrasto della criminalità

Il *predictive policing* rappresenta la metodologia più avanzata e recente in seno alle tattiche e alle strategie di polizia per il contrasto del fenomeno criminale. Affonda le proprie radici direttamente nel *problem-oriented policing* (Goldstein, 2001) e nell'*intelligent-led policing* (Ratcliffe, 2008), ovvero i due dei metodi che più hanno innovato l'attività di polizia negli ultimi quarant'anni, portandola da una risposta reattiva alle questioni criminali verso una risposta proattiva e preventiva. Il *problem-oriented policing* è un approccio di contrasto del crimine nato verso la fine degli anni '70 negli Stati Uniti ed ha mutato radicalmente l'attività delle forze dell'ordine, basandosi sulla prevenzione, su una risposta proattiva degli agenti di polizia e sulla cooperazione con altri soggetti, pubblici e privati (es. enti pubblici, comunità di cittadini) (Goldstein, 1979). Tale metodo di prevenzione è stato successivamente sintetizzato nell'acronimo S.A.R.A. (Braga, 2002):

- *scanning*, ossia identificazione di una potenziale questione criminale e definizione di questa come necessitante di un'azione di contrasto oppure no;
- *analysis*, ossia raccolta e analisi dei dati sul problema identificato, per determinarne la portata, la natura e le cause;
- *response*, ossia utilizzo delle informazioni ricavate dall'analisi per fornire una risposta adeguata al problema;
- *assessment*, ossia valutazione dell'efficacia della risposta al problema.

Il secondo approccio, ossia l'*intelligent-led policing*, rappresenta una strategia di polizia affine al *problem-oriented policing*, da cui però si differenzia per l'attenzione posta non sui problemi di fondo che causano la criminalità, ma sull'analisi delle informazioni, indicando in tale attività la base di partenza per la successiva

**La polizia predittiva è figlia del *problem-oriented policing* e dell'*intelligent-led policing*, che hanno condotto l'attività di polizia verso una risposta proattiva.**

pianificazione tattica e strategica delle forze di polizia. Questi strumenti, tecniche e processi, infatti, consentono alle autorità di pubblica sicurezza di identificare e caratterizzare le tendenze, i modelli e le relazioni nei dati associati ai comportamenti illegali e di agire conseguentemente (Ratcliffe, 2008; Beck & McCue, 2009).

Partendo da questi presupposti, il *predictive policing* ha però compiuto un ulteriore passo avanti: la peculiarità della polizia predittiva è, infatti, la sua capacità di anticipare o predire il crimine. Questo grazie ai recenti strumenti ICT (*Information and Communication Technology*), che hanno reso possibile analisi sempre più dettagliate e qualitativamente migliori della grande mole di informazioni di cui dispongono le forze dell'ordine nell'attività di contrasto della criminalità per la gestione della sicurezza a livello urbano nella società contemporanea.

**La peculiarità della polizia predittiva è, infatti, la sua capacità di anticipare o predire il crimine attraverso strumenti ICT.**

## La lotta al crimine nella società dell'informazione

Nella società contemporanea, la cosiddetta società dell'informazione (Sartori, 2012), un'infinità di dati investe, ogni giorno di più, qualunque aspetto della vita quotidiana. Questo incremento quantitativo di informazioni sta producendo un cambiamento nell'analisi dei dati (Schonberger & Cukier, 2013)<sup>2</sup>, per cui è divenuto sempre più problematico recuperare informazioni

<sup>2</sup> "Per farvi un'idea di quanto sia già avanzata la rivoluzione dell'informazione, considerate alcune tendenze che stanno investendo la società. Il nostro universo digitale si sta espandendo costantemente. Pensate all'astronomia. Nel 2000, quando è iniziata la Sloan Digital Sky Survey, nelle primissime settimane di attività il telescopio del New Mexico ha raccolto più dati di quanti ne erano stati accumulati nell'intera storia dell'astronomia. Nel 2010, l'archivio del progetto conteneva ben 140 terabyte d'informazioni. Ma un suo successore, il Large Synoptic Survey Telescope installato in Cile, che dovrebbe entrare in funzione nel 2016, acquisirà la stessa quantità di dati ogni cinque giorni". Vedi V. Schonberger – K. Cukier, Big Data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà, Milano, 2013, p. 15.

**Nella società contemporanea le forze dell'ordine devono fronteggiare il crimine dovendo maneggiare un grande numero di informazioni.**

significative. Per dirlo con le parole di Naisbitt, "we are drowning in information but starved for knowledge" (Naisbitt, 1982). Pertanto, si è reso indispensabile avvalersi del processo di KDD (*Knowledge Discovery in Database*), ovvero di quell'attività informatica che, adottando una serie di tecniche e metodologie, ricomprese nel concetto di ICT, ha per oggetto l'estrazione di conoscenza a partire dalla grande quantità di dati immagazzinati nei database (Dulli & Furini & Peron, 2009). Fulcro di tale processo sono gli algoritmi di *data mining*, grazie ai quali è possibile "esplorare" i dati e individuarne similitudini, al fine di poterli eventualmente "modellizzare".

Nell'attuale contesto storico, quindi, l'attività di *data mining* coinvolge anche le forze dell'ordine, che devono fronteggiare il crimine dovendo maneggiare un gran numero di informazioni. A tale compito, nei recenti progetti di polizia predittiva e di *predictive urban security*, come eSecurity, sono dunque deputati gli strumenti ICT. Questo nuovo paradigma di contrasto del crimine ha ricombinato in maniera del tutto innovativa, grazie all'utilizzo delle suddette tecnologie, le conoscenze provenienti dalle teorie razionali e dalla criminologia ambientale, gli schemi operativi risultanti dal *problem-oriented policing* (POP) e dall'*intelligent-led policing* e la grande quantità di dati di cui può disporre la polizia del XXI secolo. L'attuazione del paradigma del *predictive policing* ha trovato riscontro in numerose esperienze internazionali, riconducibili essenzialmente a tre progetti:

1. il software Blue CRUSH (*Crime Reduction Utilizing Statistical History*), sviluppato nel 2005 dall'IBM in collaborazione con l'Università e la polizia di Memphis in Tennessee (USA), il cui obiettivo iniziale era la diminuzione della criminalità di strada che deprimeva l'economia locale e grazie al quale è stata possibile una riduzione del 25% dei reati contro la persona e contro il patrimonio, rispetto ai mesi immediatamente antecedenti la messa in opera del programma;
2. il software PredPol (*Predictive Policing*), sviluppato nel 2011 da un team di ricercatori dell'Università di Santa Clara e dell'Università della California, in collaborazione con la polizia di Santa Cruz e Los

Angeles in California (USA), che, in un contesto economico di tagli al personale e alle risorse dei dipartimenti di polizia, ha contribuito ad una riduzione tra il 12% (a Foothill, Los Angeles) e il 25% (a Santa Cruz) dei reati contro il patrimonio;

3. il caso di Trafford (UK), coordinato dai ricercatori del Jill Dando *Institute of security and crime science* di Londra con la collaborazione della polizia di Greater Manchester (UK), il cui obiettivo era ridurre la multi-vittimizzazione in caso di furto in abitazione e grazie al quale si è riscontrata una diminuzione del 26% dei furti nelle case rispetto ai mesi antecedenti alla realizzazione del progetto.

Ciascuna delle esperienze appena menzionate sarà analizzata nella successiva sezione. L'esame tratterà inizialmente il contesto nel quale sono stati innestati i progetti; successivamente, sarà illustrato lo schema operativo concretamente posto in essere; infine, saranno presentati i risultati che ogni progetto ha conseguito.

## Concetti fondamentali

1. Il *predictive policing* è una strategia di polizia che utilizza informazioni e sviluppa analisi avanzate per la previsione delle zone a più alta densità criminale in ambito cittadino, al fine di prevenire la criminalità futura.
2. Il fondamento teorico della polizia predittiva è rappresentato dalle teorie razionali e dalla criminologia ambientale, le quali ritengono che il crimine sia prevedibile sulla base dell'assunto che un criminale deciderà di commettere un delitto ogni qual volta il crimine sia altamente desiderabile e vi sia l'opportunità di commetterlo. In particolare, tali teorie sottolineano che le attività quotidiane dei criminali e delle vittime, insieme con le caratteristiche spazio-temporali, influenzano il dove e il quando sia più probabile che avvenga un crimine.
3. Il fondamento strategico del *predictive policing* è rappresentato dal *POP* (*problem-oriented policing*) e dall'*intelligent-led policing*. Il primo ha guidato l'attività di polizia da una risposta reattiva ad una risposta proattiva e preventiva; il secondo ne ha modificato la base di partenza, indicando nell'attività di analisi delle informazioni il punto iniziale per la successiva pianificazione tattica e strategica delle forze di polizia.
4. Nella società contemporanea, la cosiddetta società dell'informazione, una grande quantità di dati investe, ogni giorno di più, qualunque aspetto della vita quotidiana, e dunque anche l'attività di polizia. La difficoltà nel reperire informazioni significative ha indotto ad utilizzare una serie di tecniche e metodologie, ricomprese nel concetto di *ICT*, che hanno per oggetto l'estrazione di conoscenza a partire dai numerosi dati immagazzinati nei *database*. Fulcro di tale processo sono gli algoritmi di *data mining*.
5. Da questo *background* deriva la possibilità, inedita per le forze dell'ordine, di interrompere il meccanismo causale del crimine, anticipandolo. Infatti, anni di dati ed informazioni sugli eventi criminali vengono raccolti e catalogati in banche dati integrate e successivamente inseriti in un *software* che li analizza. Con i dati e gli *hot spot* individuati le forze dell'ordine possono meglio posizionarsi ed esser presenti nelle zone in cui è "previsto" che si verifichi un reato quel giorno. L'operazione di polizia conduce a dei risultati che vengono inviati al *database* per essere ricompresi tra i dati e le informazioni da analizzare. E il ciclo si ripete.



# 02

## Le esperienze di polizia predittiva a livello internazionale



Nelle prossime sezioni, l'attenzione sarà posta su alcuni dei principali progetti internazionali di *predictive policing*, ovvero quelli che, allo stato attuale, hanno dimostrato di fornire i risultati più rilevanti in ambito di prevenzione, gestione e contrasto della criminalità. Di ciascuna esperienza saranno brevemente delineati il contesto in cui questa si inserisce, il suo funzionamento in concreto e gli esiti cui ha condotto.

## Il caso di Memphis in Tennessee ed il software Blue CRUSH (USA)

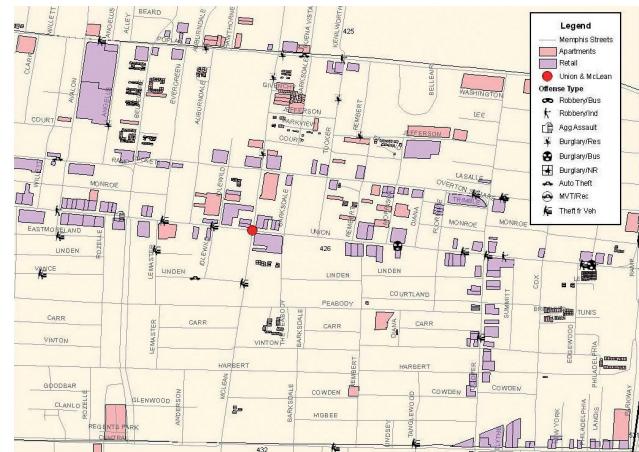
Uno dei primi progetti di polizia predittiva a livello internazionale è stato attivato nella città di Memphis in Tennessee (USA), dove nel 2005 è stato testato il software di analisi predittiva Blue C.R.U.S.H. (*Crime Reduction Utilizing Statistical History*), sviluppato da IBM (Muller, 2011).

### Il contesto

Nella città di Memphis<sup>3</sup>, nel 2005 si è verificato un considerevole incremento dei tassi di criminalità, che ha aumentato non solo il senso di insicurezza dei cittadini, ma ha anche depresso l'economia locale, tanto da spingere le autorità locali e il dipartimento di polizia della città (*Memphis Police Department - MPD*) ad elaborare una nuova strategia di contrasto e gestione della criminalità (Armstrong, 2013). In tale ottica, sono stati contattati il professor Janikowski dell'Università di Memphis e l'azienda tecnologica IBM, al fine di sviluppare un software di analisi predittiva: Blue C.R.U.S.H. (Figura 4). Il progetto pilota è partito nel 2005 ed è stato testato solo in alcuni distretti del territorio cittadino, prima di essere esteso a tutta la città (Armstrong, 2013).

<sup>3</sup> Memphis ha una popolazione di circa 677.000 abitanti, risultando così la città più popolosa dello Stato del Tennessee (USA).

Figura 4. Mappa degli hot spot creata con il software Blue C.R.U.S.H.



Fonte: IBM, [www-03.ibm.com/press/us/en/pressrelease/32169.wss](http://www-03.ibm.com/press/us/en/pressrelease/32169.wss)

### Lo schema operativo

Nella fase operativa dell'intervento di polizia predittiva, i ricercatori hanno utilizzato il software SPSS dell'IBM per l'analisi delle informazioni, ricorrendo non solo ai dati georiferiti relativi ai crimini passati e presenti, ma anche ad una serie di altri dati rilevanti ai fini del progetto (Armstrong, 2013), come variabili ambientali e socio-demografiche. Ad esempio, l'analista della polizia poteva scegliere di visualizzare, in prima battuta, una mappa con l'indicazione dei più recenti furti in appartamento georiferiti. Poi, poteva aggiungere in seconda battuta un layer costituito da una variabile socio-demografica, quale, ad esempio, quella degli indirizzi degli studenti che erano stati assenti a scuola ed, infine, aggiungere un terzo strato di dati recante l'indicazione di chi, tra quegli studenti, avesse precedenti penali per furto (Vlahos, 2012). Inoltre, i ricercatori si sono avvalsi del software ArcGIS, un sistema informativo che permette l'acquisizione, la registrazione, l'analisi, nonché la visualizzazione geografica e la restituzione di informazioni derivanti da dati presenti nel database.

### I risultati

Durante gli anni in cui la polizia di Memphis si è servita dell'analisi informatica per la prevenzione del crimine, i reati contro la persona ed il patrimonio presi in considerazione

razione nell'ambito del progetto (ad esempio, aggressione, furto di auto, furto in abitazione) sono diminuiti del 25%. A onor del vero, i risultati possono cambiare a seconda dell'anno scelto per effettuare il confronto. Nel caso di Memphis, confrontare i dati attuali con quelli del 2005 ha senso, perché è l'anno in cui è stato introdotto il software Blue CRUSH. Tale anno, però, è stato anche quello con il più alto tasso di criminalità dell'intero decennio, per cui, comparandolo con gli altri, tutti i restanti anni paiono migliori. Si può ovviare a tali problemi di comparazione confrontando i tassi di criminalità del quinquennio 2006/2010 (ovvero gli anni col software Blue CRUSH operativo) con quelli del quinquennio antecedente, ossia 2001/2005: visti così, i risultati ottenuti da Blue CRUSH a Memphis sono notevoli, ma non "miracolosi", dato che, ad esempio, i reati contro il patrimonio sono diminuiti solo dell'8% (Vlahos, 2012).

**A Los Angeles (USA) i crimini si sono ridotti del 12% con l'introduzione del software Predpol.**

## Il caso di Los Angeles in California ed il software PredPol (USA)

Una delle più importanti esperienze internazionali di polizia predittiva è stata attuata in California, dove, grazie alla collaborazione tra il locale dipartimento di polizia, l'Università di Santa Clara e l'Università della California, è stato sviluppato il software PredPol (abbreviazione di *Predictive Policing*).

### Il contesto

Dopo una prima fase di sviluppo nella cittadina di Santa Cruz (in California), dove il software predittivo ha ottenuto risultati formidabili (come, ad esempio, una riduzione del 25% dei reati contro il patrimonio nei mesi di giugno e luglio 2013 rispetto ai corrispondenti mesi del 2012), il software PredPol è stato messo a punto con un test più probante nella città di Los Angeles, sempre in California (USA)<sup>4</sup> (Friend,

2013). I ricercatori hanno testato il programma inizialmente nel distretto cittadino di Foothill, che ha una popolazione di circa 300.000 abitanti, e hanno successivamente confrontato i risultati ottenuti con quelli raggiunti mediante l'utilizzo delle *best practices* del Los Angeles Police Department (LAPD).

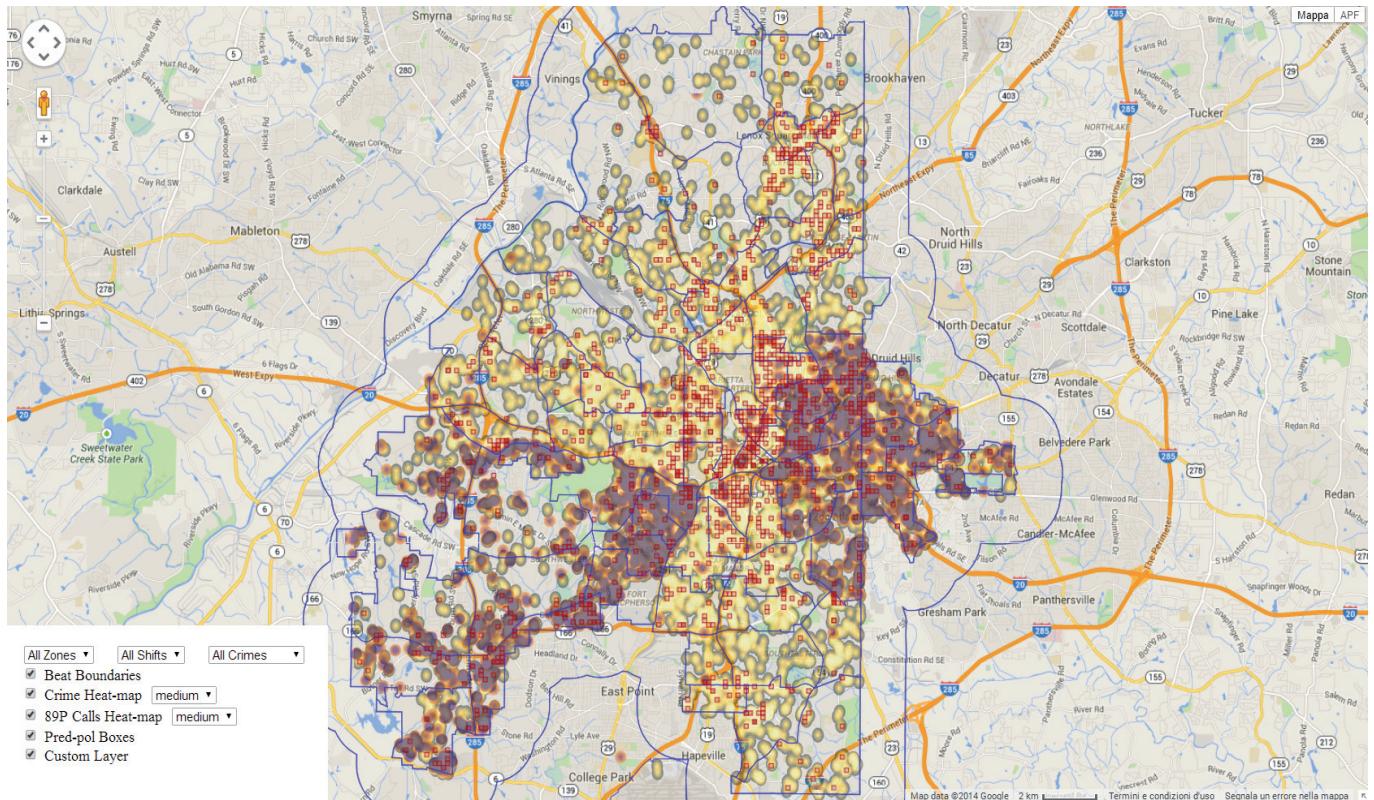
### Lo schema operativo

Il software su cui si basa il programma PredPol utilizza i dati storici relativi al tempo e al luogo in cui sono avvenuti i crimini passati presi in considerazione dal progetto (ovvero reati contro il patrimonio, come i furti in abitazione, i furti di veicoli e i furti di oggetti da veicoli). Sulla base di tali informazioni, ogni giorno vengono prodotti *report* e mappe di rischio predittivo, sulle quali sono segnate le 15 aree a più alto rischio criminale, suddivise in riquadri di 150 metri per 150 metri. In alcuni giorni della settimana, gli analisti consegnano agli agenti le mappe prodotte utilizzando i metodi tradizionali del LAPD. In altri, invece, vengono fornite le mappe disegnate utilizzando l'algoritmo predittivo presente in PredPol (Friend, 2013). Agli agenti non viene specificata la provenienza ed i metodi usati per creare le mappe: "abbiamo detto loro solo di andare negli hot spot indicati e di utilizzare le proprie conoscenze, le competenze e l'esperienza maturata per determinare gli interventi da compiere" - ha dichiarato Malinowski, capo della divisione *Real-time analysis* del LAPD (Garrett, 2013). Lo schema operativo della polizia predittiva, così come elaborato per la città di Los Angeles, si compone di diverse fasi:

1. predisposizione di un *database* mediante l'invio di una serie di *input* relativi agli eventi criminali;
2. previsione dei probabili *hot spot*, mediante l'incrocio dell'algoritmo predittivo con le informazioni sui crimini passati contenute nel *database*;
3. creazione della *heat map* con successiva produzione della mappa operativa (Figure 5-6);
4. operazione di polizia, *record* ed invio al *database* dei dati registrati sul campo;
5. ricerca dei *trend* e dei modelli criminali ad opera degli analisti.

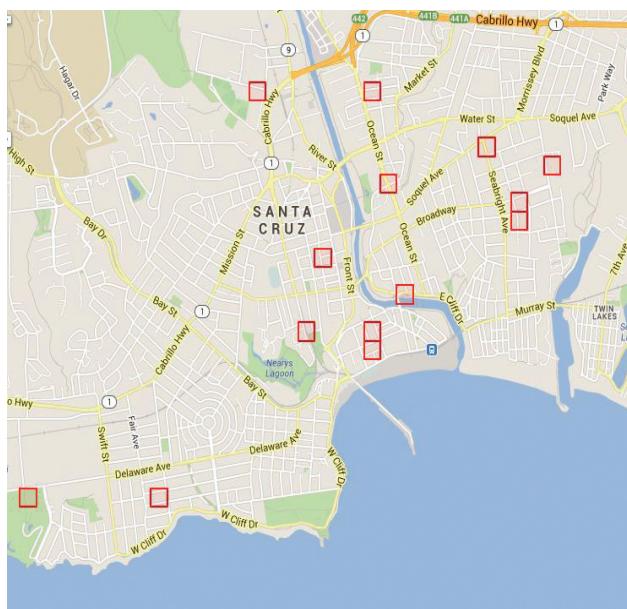
<sup>4</sup> Los Angeles ha una popolazione di circa 3.858.000 abitanti, risultando così la più grande città dello Stato della California e la seconda di tutti gli USA.

Figura 5. Heat map dei crimini elaborata con il software PredPol



Fonte: PredPol, [www.predpol.com](http://www.predpol.com)

Figura 6. Mappa operativa dei crimini elaborata con il software PredPol



Fonte: PredPol, [www.predpol.com](http://www.predpol.com)

Per la visualizzazione geografica, i ricercatori si sono avvalsi, in questo caso, delle mappe prodotte dal servizio Google Maps.

## I risultati

I risultati del test di polizia predittiva effettuato a Los Angeles sono stati sorprendenti: l'algoritmo è stato due volte più preciso delle *best practices* del LAPD. Inoltre, mentre in tutta Los Angeles i reati contro il patrimonio subivano un incremento dello 0,4%, a Foothill diminuivano del 12%. È chiaro che la pubblicazione delle statistiche criminali rappresenta sempre un “affare” politicamente sensibile, esposto al rischio concreto che i dati siano ritoccati per fare presa sull’opinione pubblica, ma in questo caso l’uso del software di predizione è stata l’unica variabile influente, non essendoci stati aumenti nel pattugliamento o cambi di turni. In seguito a tali successi, il LAPD ha ampliato il programma ad altri distretti di Los Angeles, arrivando a coprire una popolazione totale di 1,5 milioni di persone, divenendo così il più grande dipartimento di polizia a sposare la causa del *predictive policing* (Friend, 2013).

## Il caso di Trafford - Greater Manchester (UK)

Storicamente, il Regno Unito ha rappresentato un ideale ponte di collegamento tra le due sponde dell'Atlantico. Ciò è confermato anche nell'ambito della polizia predittiva, in cui un esperimento affine a quelli che si stanno sviluppando negli Stati Uniti è stato recentemente posto in essere a Trafford (UK)<sup>5</sup>.

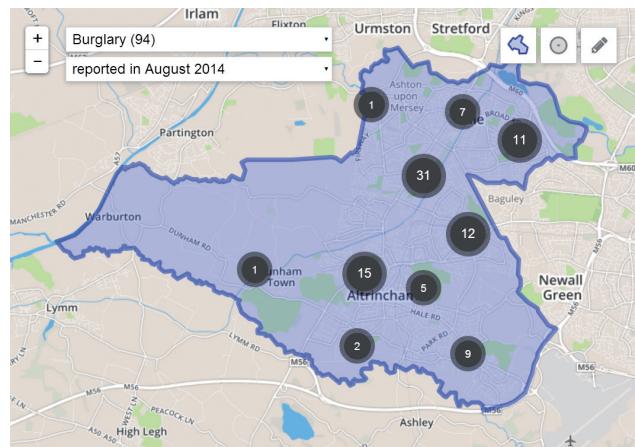
### Il contesto

Il progetto, sviluppato dai ricercatori del Jill Dando Institute of security and crime science di Londra (UK), in collaborazione con la polizia locale di Trafford, ha avuto come obiettivo iniziale la riduzione della crescente ondata di furti in abitazione (Figura 7). Il progetto ha rappresentato l'adattamento pratico di vari studi scientifici (Bowers & Johnson, 2005), che, dopo aver analizzato come la multivittimizzazione si ripeteva nel tempo e nello spazio, proponevano di collocare in maniera più efficiente, nelle settimane immediatamente successive al reato, le forze di polizia nelle zone in cui erano stati commessi furti in abitazione, al fine di impedire la *repeat victimization* (RV; "multivittimizzazione") e la *near repeat victimization* (NRV; "multivittimizzazione del vicino"). Infatti, per usare le parole di Pease, "*nei furti in abitazione il rischio di un ulteriore crimine è più alto nelle vicinanze della casa della vittima iniziale e si espande per circa 400 metri, ma scompare tra le 6 settimane e i 2 mesi. Per questo motivo, invece di mappare i furti in maniera statica, dobbiamo mappare dinamicamente il rischio per le case vicine, in modo tale che rifletta la diminuzione del rischio nel tempo*" (Jones, n.d.).

### Lo schema operativo

Il programma posto in essere a Trafford ha previsto in una fase iniziale la mappatura delle abitazioni che in precedenza avevano subito un furto e delle *buffer zones* ("zone contigue"), cui sono stati associati diversi colori, a seconda del rischio calcolato di multivittimizzazione (Jones, n.d.). Successivamente, le aree così individuate sono state prese in considerazione dalla polizia per una migliore distribuzione degli interventi e delle risorse, con l'obiettivo di potenziare la presenza di un "guardiano capace"<sup>6</sup> all'in-

Figura 7. Mappa di rischio dinamica di Trafford (UK) con indicazione delle fasce orarie dei furti in abitazione



Fonte: Home Office, [www.police.uk/greater-manchester/M1/crime/](http://www.police.uk/greater-manchester/M1/crime/)

terno delle aree più a rischio nei momenti chiave, così da scoraggiare la recidiva dei furti in abitazione.

Per l'elaborazione dei dati relativi ai furti in abitazione e per la loro visualizzazione geografica, i ricercatori si sono avvalsi del *Near repeat calculator*, un software in grado di identificare i casi statisticamente più rilevanti di multivittimizzazione e di fornire le mappe con l'indicazione spazio-temporale dei crimini avvenuti in seguito al furto in abitazione originario (Chainey, 2012).

### I risultati

Tale approccio è stato originariamente adottato per un periodo di 12 mesi nel biennio 2010/2011 ed ha ottenuto risultati che dimostrano una chiara riduzione dei furti domestici: si è riscontrata, infatti, una diminuzione del 26% dei furti in abitazione rispetto ai mesi antecedenti alla realizzazione del progetto (Jones, n.d.). Inoltre, un altro grande risultato conseguito da tale programma di *predictive policing*, considerato anche l'attuale contesto economico, è stato l'aver consentito un utilizzo più efficiente ed efficace delle scarse risorse di polizia.

<sup>5</sup> Trafford è un *metropolitan borough* di Greater Manchester in Inghilterra (UK) ed ha una popolazione di 213.000 abitanti.

<sup>6</sup> Si veda il capitolo 1. Nella teoria delle attività di routine (routine activity theory) è così definita la persona che sia in grado di impedire che un crimine sia portato a compimento. Per guardiano, gli autori di tale teoria, intendono non solo gli agenti delle forze dell'ordine, ma anche chiunque, con la sola presenza, sia in grado di fermare l'autore del probabile delitto (ad esempio, parenti, passanti, vicini) o ulteriori sistemi di controllo automatizzati (ad esempio, telecamere a circuito chiuso).

**I furti in abitazione sono diminuiti del 26% a Trafford (UK) grazie al *near repeat calculator*.**

## Concetti fondamentali

1. *Blue CRUSH.* Sviluppato nel 2005 dall'IBM, in collaborazione con l'Università e la polizia di Memphis in Tennessee (USA), il software Blue C.R.U.S.H. (*Crime Reduction Utilizing Statistical History*) ha l'obiettivo di aiutare le forze dell'ordine nel contrasto della criminalità di strada, che tende a deprimere l'economia locale e ad aumentare il senso di insicurezza dei cittadini. Il software ha il compito di analizzare i dati georiferiti relativi ai crimini passati, combinandoli con una serie di variabili ambientali e socio-demografiche, in modo da predire le future concentrazioni dei reati. Fine ultimo è la generazione di mappe e grafici dei *trend* della criminalità in base al tipo di reato, all'ora del giorno e al giorno della settimana, che possano supportare le forze dell'ordine nell'attività di prevenzione del crimine. Confrontando i tassi di criminalità del quinquennio 2006/2010 (con Blue CRUSH operativo) con quelli del quinquennio 2001/2005, le autorità locali e le forze di polizia hanno notato una diminuzione dei reati contro il patrimonio dell'8%.
2. *PredPol.* Lanciato nel 2011, il software PredPol (abbreviazione di *Predictive Policing*) è stato messo a punto nella cittadina di Santa Cruz in California (USA), prima di essere testato nel distretto di Foothill, a Los Angeles. Sviluppato da un team composto da agenti dei locali dipartimenti di polizia e da ricercatori dell'Università di Santa Clara e dell'Università della California, il software ha l'obiettivo di contribuire alla riduzione di alcuni tipi di reati contro il patrimonio (come ad esempio, il furto in abitazione, il furto di auto ed il furto di oggetti dal veicolo). Per raggiungere questo scopo, ogni giorno PredPol analizza i dati storici relativi al tempo e al luogo in cui sono avvenuti i crimini passati ed elabora mappe di rischio, che il dipartimento di polizia interessato provvede a distribuire agli agenti solo in alcuni giorni della settimana. Negli altri, invece, gli analisti consegnano agli agenti mappe non predittive, prodotte utilizzando le *best practices* del LAPD. Agli agenti non viene specificata la provenienza ed i metodi usati per creare le mappe, così da poter confrontare i risultati ottenuti con l'una e con l'altra metodologia. I risultati del test di polizia predittiva effettuato a Los Angeles, sono stati sorprendenti: il software PredPol è stato due volte più preciso delle "best practices" del LAPD. Inoltre, mentre in tutta Los Angeles i reati contro il patrimonio subivano un incremento dello 0,4%, a Foothill diminuivano del 12%.
3. *Near repeat calculator.* Il progetto posto in essere a Trafford (UK), sviluppato dai ricercatori del Jill Dando Institute of security and crime science di Londra (UK) in collaborazione con la polizia locale, ha come obiettivo la riduzione della crescente ondata di furti in abitazione in Inghilterra. Per raggiungerlo, i ricercatori hanno inizialmente mappato le abitazioni della città che avevano subito in precedenza un furto e le *buffer zones* ("zone contigue"), cui sono stati associati diversi colori, a seconda del rischio di multivittimizzazione. Per l'elaborazione dei dati relativi ai furti in abitazione e per la loro visualizzazione geografica, i ricercatori si sono avvalsi del *Near repeat calculator*, un software in grado di identificare i casi statisticamente più rilevanti di multivittimizzazione e di fornire mappe dinamiche con l'indicazione spazio-temporale dei crimini avvenuti in seguito al furto in abitazione originario. Questo intervento ha comportato una riduzione del 26% dei furti in abitazione rispetto ai mesi antecedenti alla realizzazione del progetto.

# 03

## Metodi statistici per la previsione della concentrazione della criminalità

A chalkboard featuring several mathematical expressions and arrows. The expressions include:  
Top row:  $\sqrt{9}$ ,  $\frac{9}{9}$ ,  $\frac{9+9}{9}$   
Bottom row:  $\sqrt{9}-\frac{9}{9}$ ,  $\sqrt{9}+\frac{9}{9}$   
Center row:  $\sqrt{9}+9-$ ,  $9-\frac{9}{9}$ ,  $\sqrt{9!}-\frac{9}{9}$   
An upward-pointing arrow is positioned between the top and bottom rows. A horizontal double-headed arrow is located below the center row, pointing from left to right.

$$\sqrt{9}$$
$$\frac{9}{9}$$
$$\frac{9+9}{9}$$
$$\sqrt{9}-\frac{9}{9}$$
$$\sqrt{9}+\frac{9}{9}$$
$$\sqrt{9}+9-$$
$$9-\frac{9}{9}$$
$$\sqrt{9!}-\frac{9}{9}$$

Tutti i recenti progetti di polizia predittiva, volti a prevenire il verificarsi di un episodio criminale futuro, fanno uso di dati sui crimini passati (e, in alcuni casi, di una serie ulteriore di informazioni) che, grazie anche ai nuovi strumenti ICT, possono condurre alla scoperta di tendenze e modelli predittivi del comportamento criminale. La capacità del *predictive policing* di mettere a disposizione informazioni qualitativamente migliori nell'ambito dell'attività di polizia può consentire alle forze dell'ordine di contrastare il crimine più efficacemente, giocando d'anticipo. Infatti, gli agenti di polizia possono meglio posizionare le proprie pattuglie ed organizzare i propri interventi, in modo tale d'essere presenti e da poter condurre gli interventi di polizia in maniera mirata nelle zone in cui è previsto che si verifichi un reato quel giorno.

Allo stato attuale, i progetti di polizia predittiva utilizzano differenti metodologie a seconda dell'obiettivo che si vuole raggiungere e del *background* di informazioni di cui si dispone. In particolare, i metodi utilizzati possono essere riassunti in tre categorie, che ricoprono specifici modelli di analisi:

1. l'*hot spot analysis*, il *data mining*, la regressione statistica e il metodo *near-repeat* vengono generalmente utilizzati per identificare dove è più probabile che avvenga un reato durante uno specifico arco temporale;
2. i *modelli spazio-temporali* sono impiegati per individuare quando è verosimile che avvenga un crimine;
3. la *risk terrain analysis* è, invece, utile se si vogliono conoscere i fattori geospaziali che aumentano il rischio che avvenga un evento criminale.

**Tabella 1. Classi di tecniche predittive**

<b>Analytic Category and Primary Application</b>	<b>Predictive Technique</b>	<b>Class</b>			
		<b>Classical</b>	<b>Simple</b>	<b>Complex</b>	<b>Tailored</b>
Hot spot analysis (where, using crime data only)	Grid mapping	X			X
	Covering ellipses	X			
	Kernel density	X			
	heuristics		X		X
Regression methods (where, using a range of data)	Linear	X	X		
	Stepwise	X		X	
	Splines			X	X
	Leading indicators	X			X
Data mining (where, using a range of data)	Clustering	X		X	
	Classification	X		X	
Near-repeat (where, over next few days, using crime data only)	Self-exciting point process			X	
	ProMap			X	
	heuristic		X		
Spatiotemporal analysis (when, using crime and temporal data)	heat maps	X	X		X
	Additive model			X	
	Seasonality	X			
Risk terrain analysis (where, using geography associated with risk)	Geospatial predictive analysis			X	X
	Risk terrain modeling		X		X

Fonte: RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013, p. 19

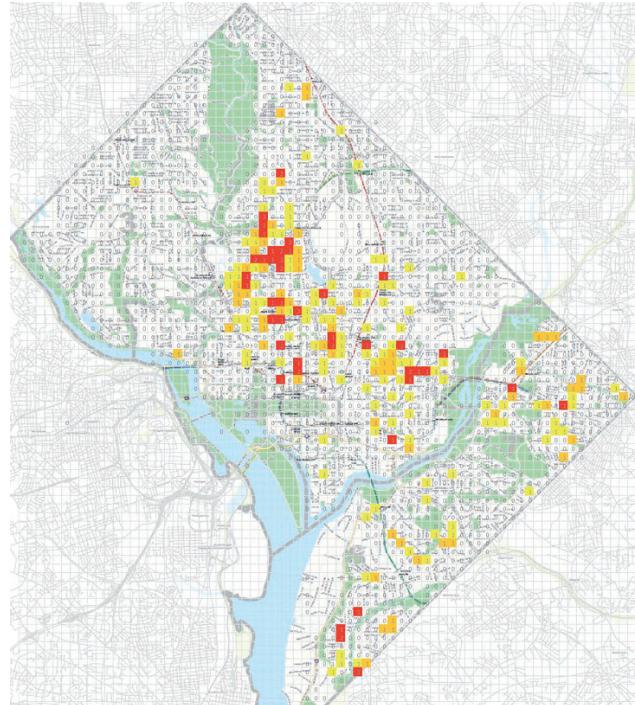
Nei prossimi paragrafi, saranno tratteggiate le tecniche predittive maggiormente adoperate nei progetti di *predictive policing* e di *predictive urban security* (Figura 7), per la previsione di dove e quando sia più probabile che venga posto in essere un comportamento criminale (RAND, 2013).

## Hot spot analysis

Per *hot spot analysis*, si intendono un insieme di tecniche statistiche che mirano alla predizione delle aree nelle quali, sulla base dei dati storici dei crimini, la concretizzazione di un reato può essere più frequente (le cosiddette “zone calde” appunto). In criminologia, tali luoghi vengono identificati col nome di *hot spot* (o anche *hot dots*, *hot areas*, *hot lines*, a seconda delle caratteristiche fisiche del luogo), ovvero come “quei luoghi in cui il verificarsi di un crimine è così frequente che si può definire come altamente prevedibile” (Sherman, 1995; Eck et al., 2005). Fra le tecniche statistiche più utilizzate in tale ambito possono annoverarsi:

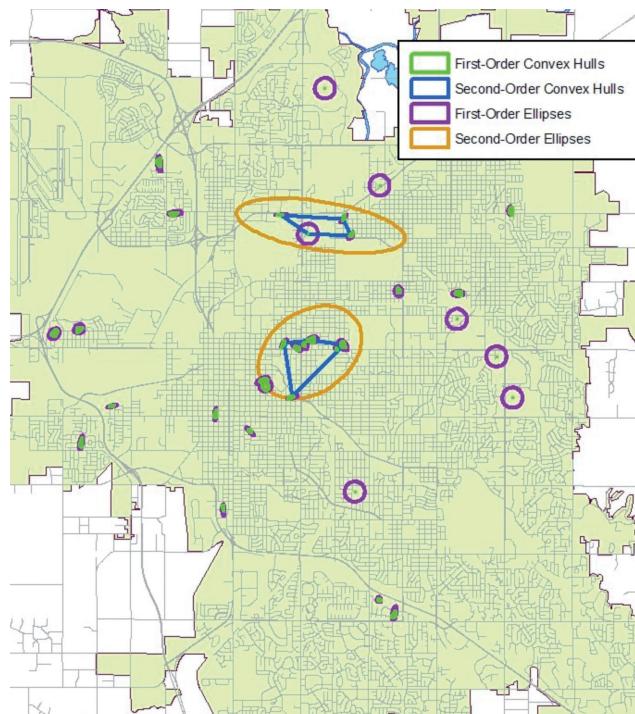
1. il *grid mapping* (Figura 8). Tale metodologia mira a suddividere, attraverso l’uso di coordinate cartesiane, l’area oggetto di analisi (ad esempio una città) in tante piccole partizioni (celle) e a misurare la quantità di reati accaduti per ogni cella. A partire dalla frequenza stimata e dalle variazioni nel numero di reati avvenuti, i risultati vengono “pesati” sulla base dei crimini avvenuti durante un determinato arco temporale (ad esempio, 6 mesi). In questo modo vengono identificati gli *hot spot*, le “zone calde”, che possono a loro volta concentrarsi in determinate aree (ad esempio, il centro città per determinati reati come il borseggio). Il sistema di analisi più utilizzato in tale ambito è il GIS (*Geographical Information System*) e tutti i software di questo derivati: ArcGIS, R e altri (Chainey e Thompson, 2012; RAND, 2013).
2. le *covering ellipses* (Figura 9). Attraverso l’identificazione degli *hot spot*, è possibile “coprire” l’area analizzata con delle ellissi, le quali inglobano *cluster* (insiemi) di vari *hot spot* vicini. Trattandosi di ellissi, queste vanno chiaramente ad inglobare anche aree non *hot* che si trovano però nelle vicinanze di gruppi di aree calde. Con l’ausilio dei moderni software è possibile “tagliare” le zone non *hot* incluse nelle ellissi, fino a giungere ad aree ristrette che includano solamente *hot spot*. È il caso del cosiddetto *two step approach*, il quale utilizza le ellissi per raggruppare geograficamente i crimini e quindi definisce gli *hot spot* come le più piccole aree comprendenti tutti i crimini di un *cluster*. Un esempio della metodologia *covering ellipse* è il *nearest neighbor hierarchical clustering* (NNH), che identifica gruppi di eventi vicini spazialmente e li raggruppa sulla base di dati criteri all’occorrenza stabiliti.

Figura 8. Grid map delle rapine a Washington D.C. (USA)



Fonte: Preliminary findings from an RTI International, Structures Decisions Corporation, and RAND project funded by NIJ OST, 2013

Figura 9. Covering ellipses create utilizzando la NNH delle rapine a Washington D.C. (USA)



Fonte: S.C. Smith and C.W. Bruce, *CrimeStat III User Workbook*, Washington, 2010, p. 56, Figure 5-10

I modelli di *clustering ellittici*, però, soffrono di due problemi fondamentali:

- a. in primo luogo, richiedono che il numero di ellissi sia definito in anticipo, ma non sempre è possibile determinarlo con precisione. Tale problema può essere ovviato empiricamente, effettuando di volta in volta *test* con diversi numeri di ellissi fino a trovare il risultato migliore;
  - b. in secondo luogo, i risultati di ogni osservazione possono essere molto sensibili ai dati in *input*. Questo problema si palesa soprattutto in riferimento ad analisi stagionali: ad esempio, se si prendono in considerazione gli effetti della variazione sui 3 mesi, le differenze possono essere molto basse; viceversa se si analizzano su 6 mesi, le differenze possono essere talmente ampie da essere fuorvianti (RAND, 2013);
3. la *Kernel density estimation* (KDE). Con tale metodo statistico, si identificano gli *hot spot* futuri sulla base della loro prossimità ai crimini attuali, facendo ricorso ad algoritmi matematici (chiamati metodi Kernel). Un Kernel è una funzione standardizzata pesata utilizzata, in questo caso, per “lisciare” (ossia dare una forma continua) i dati sull’incidenza dei crimini. In concreto, si seleziona una funzione Kernel ed un raggio d’azione (“intervallo”) e si posiziona ogni valore nel punto geografico di occorrenza di ciascun crimine. L’area attorno al punto in cui si è verificato il comportamento illecito viene “pesata”

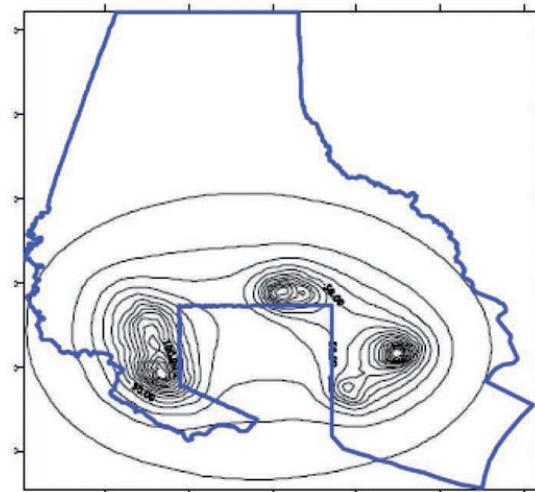
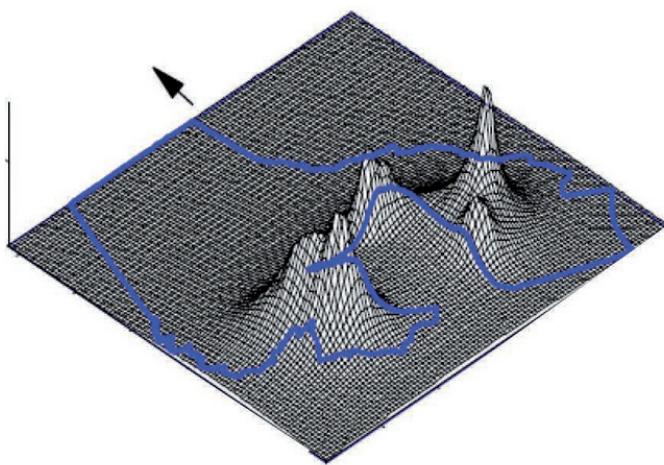
secondo la funzione Kernel (Sullivan e Perry, 2004). Il KDE (Figura 10) è, quindi, un approccio metodologico utilizzato per interpolare una superficie continua di dati, basandosi sulla localizzazione dei crimini correnti (RAND, 2013).

Esistono due tipologie di *Kernel density estimation*:

- a. la *single* KDE, che stima gli *hot spot* usando una sola variabile, quella dei crimini già avvenuti;
- b. la *dual* KDE, che invece utilizza due variabili, quella relativa ai crimini e quella concernente la densità della popolazione dell’area analizzata. Se infatti non si tiene conto della densità della popolazione, ad esempio negli studi sulla criminalità urbana, il centro città risulta essere più vittimizzato; viceversa se la si considera, le aree maggiormente a rischio risultano essere i sobborghi più disagiati della città.

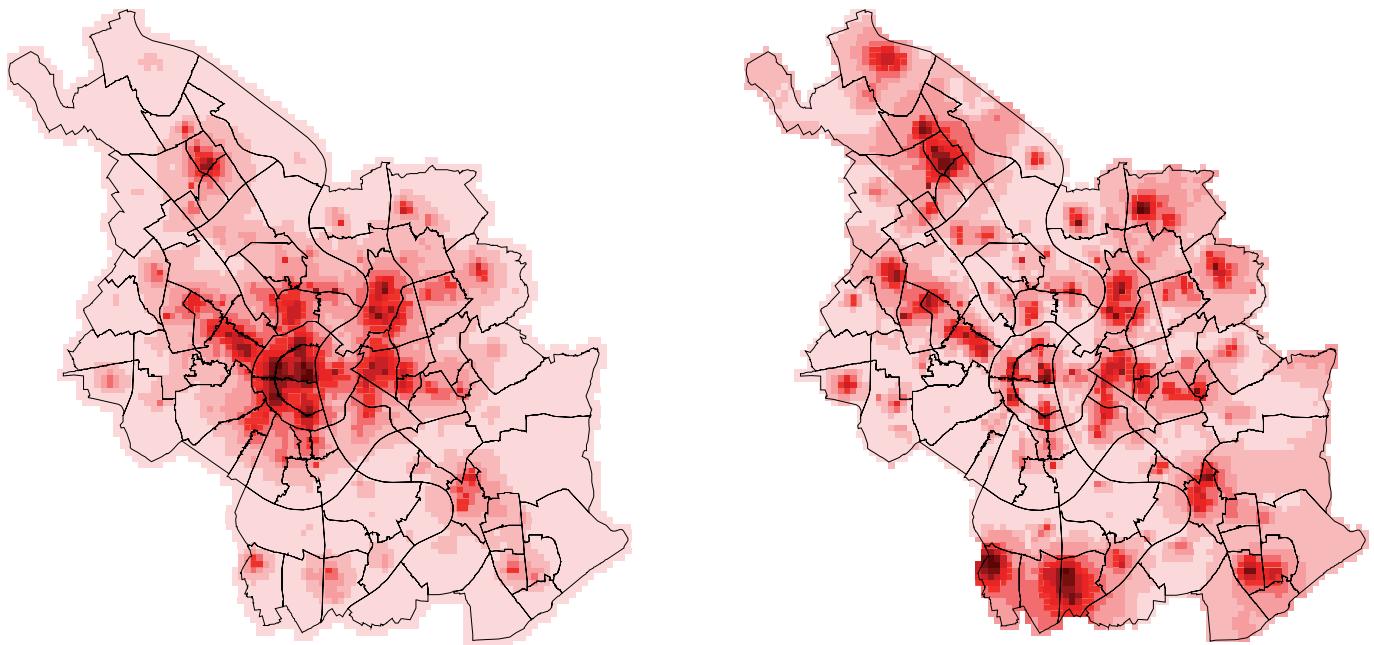
Il KDE risulta essere un metodo robusto, in cui piccoli cambiamenti nella definizione della funzione portano a piccole variazioni nei risultati. Tuttavia, non partendo da una definizione iniziale di quanti *hot spot* dovrebbero esserci, il KDE presenta un limite, relativo alla sensibilità dei risultati nei confronti dei dati di *input*, che può essere risolto solo pesando le funzioni Kernel sulla base della rilevanza del periodo di predizione. Il KDE è considerato comunque uno dei metodi di predizione più accurati (Figura 11);

Figura 10. KDE delle rapine nella contea di Baltimora in Maryland (USA) nel periodo 1996-1997



Fonte: N. Levine et al., *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations*, prepared for the NIJ, November 2004, p. 8-18, Figure 8-9

Figura 11. Confronto tra un *single KDE* e un *dual KDE* in relazione alle aggressioni avvenute Colonia (Germania) nel 2012



Fonte: D. Oberwittler e M. Wiesenhütter, "The Risk of Violent Incidents Relative to Population Density in Cologne Using the Dual Kernel Density Routine," in N. Levine and Associates, 2004, p. 8-36

4. i *metodi euristici*. In tale categoria, rientrano tutte quelle metodologie basate sull'esperienza e sulle capacità professionali degli agenti delle forze dell'ordine. Vi sono ricomprese (Eck et al., 2005; RAND, 2013):

a. l'*identificazione manuale* degli *hot spot* sulle *pin maps*. È un metodo basato sulla annotazione manuale degli *hot spot* sulle mappe geografiche ad opera degli agenti di polizia, che basano l'identificazione sull'esperienza dei crimini avvenuti in una determinata zona. Tale metodo è molto utilizzato, tuttavia presenta alcune problematiche: ad esempio, se l'analisi degli *hot spot* viene operata solo sulla base dei reati avvenuti in un dato punto, si rischia di visualizzare i punti in cui un crimine si reitera non come un *cluster* di crimini, ma come un punto di singolo reato, perdendo così la visione d'insieme;

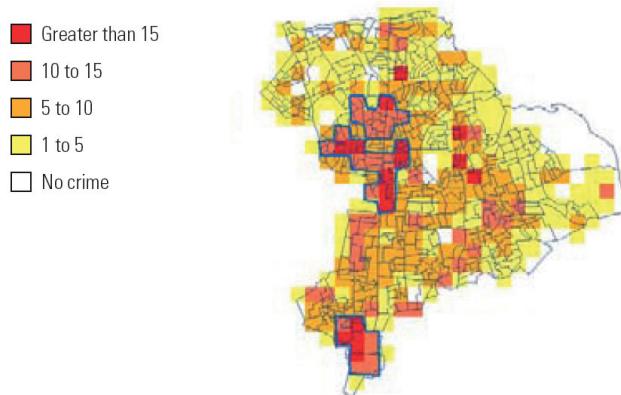
b. il *quadrat thematic mapping* (Figura 12). Consiste nella versione manuale del *grid mapping*. Questa tecnica si basa sui dati aggregati relativi ai crimini avvenuti e li visualizza come una matrice di poligoni (celle) di uguale grandezza. Anche per questo metodo ogni cella identifica le aree *hot*, che saranno di colori differenti in ragione del numero di crimini ivi registrato. Tale metodo è incluso fra quelli euristici, poiché l'intensità

di colore assegnata ad ogni cella è tipicamente determinata sulla base dell'esperienza passata dei crimini avvenuti;

c. lo *jurisdiction-bounded areas* (Figura 13). Tale metodo consiste nell'utilizzo di mappe tematiche basate su dati geografici giurisdizionali (come ad esempio i distretti, il CAP, le sezioni di censimento), in cui vengono utilizzati dei poligoni per determinare gli *hot spot*. I poligoni vengono evidenziati con una scala di colori, per cui più un poligono è scuro più tale area è considerata "a rischio". L'impiego di questo approccio può comportare l'uso del MAUP (*modifiable areal unit problem*), ovvero l'utilizzo di confini politici e/o amministrativi che riflettano la geografia dei poligoni e non realmente la struttura delle aree più a rischio. A causa di questo, spesso la selezione degli *hot spot* può riflettere la geografia amministrativa o fisica di un dato luogo, discordando quindi da un reale confine di rischio.

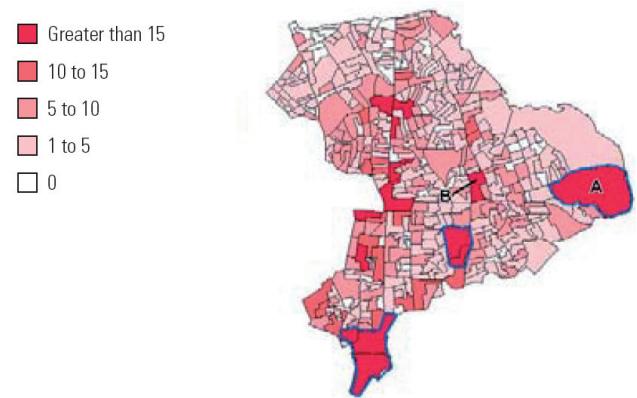
**L'*hot spot analysis* mira alla predizione delle aree nelle quali un reato può essere più frequente**

**Figura 12.** Quadrat thematic mapping per il reato furti di veicoli



Fonte: Eck et al., 2005, p.25

**Figura 13.** Mappa tematica sull'incidenza dei crimini, costruita utilizzando i confini geografici giurisdizionali per il reato furti di veicoli



Fonte: Eck et al., 2005, p.24

## Data mining

La costruzione di modelli matematici finalizzata ad effettuare previsioni viene spesso identificata con l'appellativo di "tecniche di *data mining*". Mediante tali metodologie si mira ad individuare all'interno di grandi set di dati l'esistenza di *pattern* e *trend* che siano utili agli scopi scientifici che una ricerca si propone. Nell'ambito delle tecniche di *data mining*, si distinguono (Chamont e Pin-Shuo, 2008; de Bruin et al., 2006):

1. i metodi di *regressione*, con i quali si mira a stimare il rischio che si verifichi un crimine in futuro, attraverso la relazione tra la variabile da prevedere, ossia il pericolo che venga posto in essere un comportamento illecito, con un certo numero di variabili esplicative. Questi metodi saranno trattati in dettaglio nel prosieguo (vedi p. 24);
2. i metodi di *classificazione*. Si occupano di predire una categoria di dati sulla base di un determinato *outcome*. Gli algoritmi di classificazione lavorano preliminarmente sui dati per individuare eventuali *pattern* che vanno a determinare le categorie ("classi") osservate; in seguito, questi *pattern* vengono utilizzati al fine di effettuare previsioni circa future osservazioni. Gli algoritmi di classificazione possono essere relativamente semplici, come gli alberi di decisione, o matematicamente complessi, come le reti neurali o i *support vector machine*, i quali rientrano fra i metodi cosiddetti *black box*, poiché combinano tutte le variabili servendosi di formule molto complicate e difficilmente osservabili;
3. i metodi di *clustering*. Queste tecniche suddividono le osservazioni in gruppi a seconda della loro

somiglianza in termini matematici. Questi modelli si occupano di operare previsioni partendo dall'assunto che una situazione futura sarà simile ad un *cluster* di situazioni passate. Il *clustering* può funzionare anche in presenza di un'ampia disponibilità di informazioni circa i dati oggetto di analisi. La tecnica di *clustering* da utilizzare dipende dalla natura dei dati (categorici o numerici), ma molti moderni algoritmi selezionano automaticamente la tecnica appropriata, sulla base dei dati in *input*. Infatti, gli algoritmi di *clustering* possono essere utilizzati anche nella fase di esplorazione dei dati così da rintracciare una comunanza fra i crimini. Un esempio di algoritmi di questo tipo sono quelli di tipo spaziale usati per i dati sui crimini georiferiti, al fine di rintracciare *hot spot* statisticamente significativi. Rientrano in quest'ambito gli indicatori di associazione spaziale usati per valutare l'esistenza di *cluster* spaziali di una data variabile. Uno di questi è la classe di indicatori locali di associazione spaziale (LISA), tra i quali è possibile annoverare il *Getis Ord G\* Statistic*, che compara valori locali a valori globali per determinare l'entità del *clustering* spaziale. L'*output* della *G\* Statistic* è uno z-score che indica la magnitudo ("importanza") del *cluster*. Livelli alti o bassi di z-score riflettono un elevato livello di *clustering*, mentre valori prossimi allo 0 indicano casualità spaziale;

4. i metodi *ensamble*. Tali metodi, spesso definiti come una sottocategoria dei metodi di classificazione, utilizzano semplici modelli predittivi e li combinano per avere una predizione finale. I metodi *ensamble* generano un ampio set di semplici modelli predittivi, che si combinano tra loro in diversi modi. Tra i più diffusi ci sono:

- a. i metodi *boosting*. Questi partono da un semplice modello di classificazione e iterativamente aggiungono ad esso altri semplici modelli di classificazione. L'idea è di raggiungere una maggiore accuratezza combinando più modelli semplici e cercando, dunque, un modello combinato che possa classificare in maniera corretta le occorrenze che vengono classificate scorrettamente durante l'iterazione precedente o il modello semplice di partenza. La predizione finale sarà, quindi, una media pesata di tutti i modelli di classificazione utilizzati;
- b. i metodi *random forest*. Questa classe di metodi genera iterativamente un ampio set di alberi di decisione semplici. Ad ogni iterazione, gli alberi vengono "arricchiti" con sottoinsiemi casuali di variabili e dati di *input*. La predizione finale sarà il risultato delle decisioni prese per ogni singolo albero;
- c. i metodi *bagging*. Tale tecnica genera un ampio numero di classificatori, costruendo un modello di classificazione ad ogni iterazione dell'algoritmo composto da un sottoinsieme casuale di dati di *input*. I classificatori utilizzati sono alberi di decisione e reti neurali e la predizione finale sarà una media dei risultati ottenuti (Shyam, 2006; Smith e Bruce, 2010).

Nel *data mining* riveste una parte importante la fase di valutazione, ovvero la verifica dell'accuratezza di queste metodologie, in cui ci si basa su vari indicatori, fra cui il tempo di calcolo (dato molto importante nel processo di decisione su quale metodo sia da preferire). Tale accorgimento è ancora più importante se ci sono molte variabili di *input* e, quindi, molti modi per combinarle. Durante tale processo di verifica, si ricorre alla divisione dei dati in due gruppi: uno di *training*, su cui si "costruisce" il modello, e uno di *testing*, su cui si "verifica" la correttezza dei modelli scelti. Nello specifico, si modellano i dati del set di *training* con il metodo scelto; a questo punto, si prova il metodo prescelto sul set di *testing*, così da valutare le eventuali differenze (Chen et al., 2004; Chung-Hsien et al., 2011).

**Attraverso le tecniche di *data mining*, si mira ad individuare all'interno di grandi set di dati l'esistenza di pattern e trend che siano utili agli scopi della ricerca.**

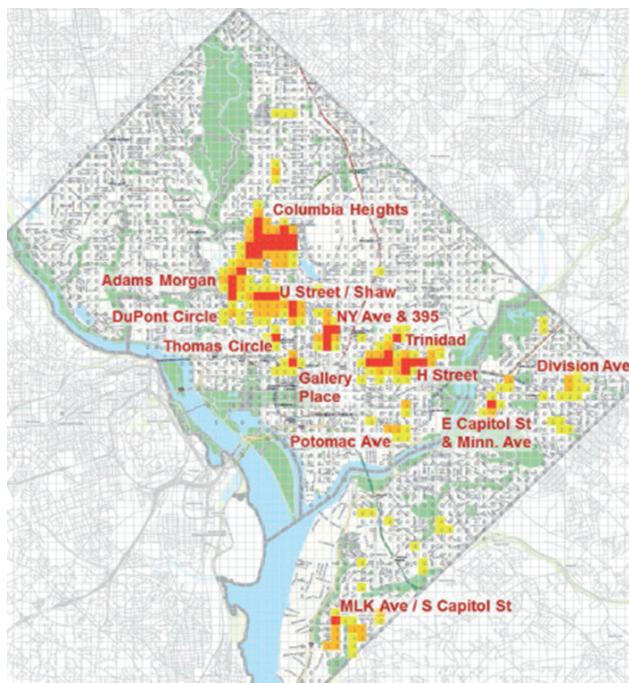
## Metodi di regressione

I metodi di regressione sono ampiamente utilizzati per la predizione della concentrazione della criminalità. Mediante la loro applicazione, si punta a stimare il rischio che si verifichi un crimine in futuro. Ciò è possibile se si mette in relazione la variabile da prevedere, ossia il pericolo che venga posto in essere un comportamento illecito, con un certo numero di variabili esplicative costituite dai dati relativi ai crimini passati e da un ampio *range* di informazioni ulteriori (Figura 14). Ad esempio, un modello di regressione per i furti in abitazione include come variabili di *input*, oltre ai furti stessi, anche il numero di altri tipi di reato o di altri episodi di disordine urbano verificatisi nel quartiere. Le tecniche di regressione sono oggi usate per l'analisi di molti tipi di crimine; tuttavia non sono esenti da problemi, che possono sorgere qualora il set di dati sia limitato o vi sia volatilità nei dati.

Alcune tecniche di regressione possono essere utilizzate per la polizia predittiva. I metodi di regressione si distinguono sulla base della relazione che intendono stimare e possono suddividersi in (Jianqiang et al., 2005):

1. *regressione di tipo lineare*. Tale modello mette in relazione due o più variabili, attraverso l'uso di un'equazione lineare. Questa relazione può essere definita come una media pesata delle variabili di *input* ed è definita lineare poiché, geometricamente, può essere rappresentata come una linea o un piano che mette in relazione le variabili di *input* o di *output*. Fra queste metodologie la più utilizzata è quella dei *least squares* ("minimi quadrati"), la quale punta ad identificare la formula che meglio si adatta ai dati osservati, minimizzando la somma dei quadrati delle deviazioni che i dati osservati hanno rispetto ai dati predetti;
2. *regressione di tipo non lineare*. Regressioni di questo tipo si avvalgono di formule matematiche più complesse di una semplice media pesata, al fine di mettere in relazione i dati di *input* con le variabili di regressione;
3. *regressione di tipo spline*. Le regressioni *spline* possono essere definite come regressioni "miste", in cui vengono utilizzati diversi metodi allo scopo di modellare i dati, variando a seconda delle differenti regioni delle variabili dipendenti. In particolare, vengono utilizzate regressioni "a tratti" (ovvero viene modellata una regressione in ogni segmento) su tutto il set delle variabili dipendenti.

Figura 14. Le rapine a Washington D.C. analizzate con la regressione multipla



Fonte: *Preliminary findings from an RTI International, Structures Decisions Corporation, and RAND project funded by NIJ OST, 2013*

Un argomento rilevante, presente in tutti i modelli di regressione, riguarda la selezione delle variabili da mettere in relazione. A tale scopo, si fa uso di diversi metodi, quali:

4. **l'esperimento empirico e la correlazione euristica.**  
In questo caso, viene calcolata la correlazione tra le variabili di *input* e di *output* e, successivamente, vengono filtrate le variabili di *input* troppo poco o eccessivamente correlate con quelle di *output*. I criteri per stabilire il filtraggio vengono definiti attraverso esperimenti e simulazioni, allo scopo di trovare un modello con un buon grado di accuratezza predittiva e per il quale ogni variabile di *input* scelta contribuisca in maniera significativa;
5. **le regressioni forward/stepwise.** Questi metodi sono euristicamente misurati in cui costruiscono modelli di regressione in maniera iterativa. Per quel che riguarda i metodi *forward*, ad ogni iterazione viene aggiunta una variabile di *input* e si procede finché non si raggiunge il risultato migliore in termini di accuratezza del modello. I metodi *stepwise*, invece, partono da modelli comprendenti tutte le variabili e procedono con l'eliminazione graduale delle variabili fino a giungere al modello più accurato. Tali metodologie hanno come maggior critica quella di essere approssimative e di fornire sovente false impressioni circa i risultati, fortemente influenzati dal-

le variabili in *input*. Quando si utilizzano regressioni *forward* possono sorgere problemi relativi all'ordine con cui vengono inserite le variabili, mentre solitamente con le regressioni *stepwise* si raggiungono buoni livelli di accuratezza predittiva;

6. **l'ottimizzazione numerica.** Questa metodologia si utilizza ponendo delle penalità sul numero di variabili incluse nella relazione, al fine solito di ottenere il miglior modello in termini di accuratezza. Ciò si raggiunge servendosi di complesse formulazioni matematiche. Fra i più usati troviamo la *least angle regression*, la *lasso regression*, l'*elastic net*.

Nella costruzione dei modelli di regressione, vengono spesso prese in considerazione variabili esterne, che possono influenzare la previsione dei crimini, come, ad esempio, l'aumento della temperatura o il miglioramento delle condizioni di vita di un'area depressa, le quali possono effettivamente avere un impatto sui crimini futuri. Tali variabili hanno come fattore positivo la facilità nell'essere identificate e sono provviste di un forte potere informativo. Un esempio è quello riportato da Bushman et al., i quali dimostrano il grado di correlazione fra l'aumento delle temperature e l'aumento di crimini violenti, anche nel caso di diverse stagioni o momenti della giornata (Bushman et al., 2005). Tali variabili, in particolare nei modelli più semplici come la regressione lineare, vengono utilizzate come coefficienti di *smoothing*, ossia quei casi in cui viene dato maggior peso agli eventi più recenti rispetto a quelli più risalenti nel tempo, al fine di incrementare il grado di affidabilità nella previsione (Jianqiang et al., 2005).

Un discorso a parte meritano i modelli di regressione spaziale, comunemente utilizzati in molti ambiti di analisi e ricerca fin dalla loro introduzione (Bernasco e Elffers, 2010). Queste tecniche vengono utilizzate in criminologia per esaminare l'influenza di particolari variabili su differenti tipi di attività criminale e offrono supporto a diverse teorie criminologiche (Anselin e Hudak, 1992; Messner e Anselin, 2004; Porter e Purser, 2010). Ad esempio, in Baller et al. (2001), il modello di regressione spaziale viene utilizzato per esaminare le variabili esplicative presenti dietro il tasso di omicidi a livello provinciale. Questo approccio lascia scoperto il *clustering* spaziale degli omicidi, così come le differenze regionali presenti nell'impatto delle variabili esplicative sul tasso di omicidi.

**Nei modelli di regressione vengono considerate variabili esterne, che possono influenzare la previsione dei crimini futuri.**

## Near repeat methods

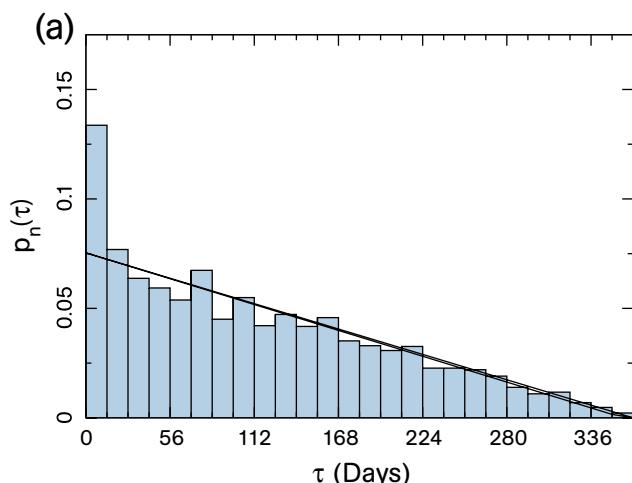
I metodi di analisi del tipo *near-repeat*, applicati alla predizione della criminalità, operano previsioni partendo dall'assunto che determinati crimini futuri si ripeteranno in zone molto vicine a quelle in cui sono avvenuti crimini in un periodo passato. Ad esempio, i furti in abitazione si ripetono sovente nelle stesse zone poiché i ladri ne conoscono le debolezze, come ad esempio lo scarso controllo da parte delle forze di polizia o la molteplicità delle vie di fuga (Mohler et al., 2011). Alcuni criminologi sostengono che i crimini si propagano negli ambienti urbani allo stesso modo di una malattia contagiosa: se in una determinata zona c'è stato un crimine, la probabilità che quel tipo di crimine si ripeta in quella stessa zona in un breve periodo di tempo aumenta (Mohler et al., 2011; Townsley et al., 2000).

Partendo da tale teoria, Mohler et al. hanno ideato un algoritmo chiamato *self-exciting process* o *earthquake modeling*. Nonostante si serva di complesse formulazioni matematiche, l'approccio di questo algoritmo è semplice: viene disegnato un reticolo sul territorio analizzato, come nel *grid mapping*, e viene poi stimato il tasso ("background rate") al cui raggiungimento è più probabile che si verifichi un nuovo crimine. Il valore di questo tasso dipende dalle caratteristiche proprie di ciascuna partizione del reticolo e dagli effetti temporali nelle zone limitrofe a quella presa in considerazione. Infatti, quando si verifica un crimine, si assume che il tasso di ripetizione di quel crimine salirà con il passare del tempo ("aftershock rate"). Se non dovesse verificarsi un nuovo crimine, questo tasso decrescerà naturalmente (Figura 15).

Questo algoritmo funziona in maniera simile a quello utilizzato per la previsione dei terremoti futuri (da cui l'appellativo *earthquake modeling*), in cui la probabilità del verificarsi di un nuovo terremoto cresce appena dopo un terremoto già avvenuto, per poi ridiscendere se il terremoto non si verifica. Per la stima dei *background rate* e degli *aftershock rate*, si ricorre all'utilizzo di simulazioni con tecnica MonteCarlo (Short et al., 2008; 2009).

**I metodi di tipo *near-repeat* operano previsioni partendo dall'assunto che i crimini futuri si ripeteranno in zone vicine a quelle in cui sono avvenuti crimini in passato.**

Figura 15. Furti in appartamento a Long Beach (California) analizzati con metodologia Near-repeat per unità abitative distanti 100 metri



Fonte: M. B. Short et al., "Measuring and Modeling Repeat and Near-Repeat Burglary Effects", *Journal of Quantitative Criminology*, 25/3: 325-339, 2009

## Analisi spazio-temporale

L'analisi spazio-temporale (*spatio-temporal analysis*) mira ad analizzare i crimini e la loro evoluzione nello spazio e nel tempo. Le metodologie utilizzate per questo tipo di analisi espandono il problema della previsione degli *hot spot*, includendo ulteriori informazioni relative alle caratteristiche ambientali e temporali dell'area di occorrenza dei crimini. L'idea di fondo è, infatti, quella di utilizzare questi dati per predire il luogo e il tempo dei futuri reati. L'analisi spazio-temporale, nella sua applicazione di base, può essere realizzata ricorrendo a semplici metodi euristici. Generalmente questi consistono in un procedimento "manuale", eseguito utilizzando interfacce grafiche essenziali. Spesso, però, si ricorre a metodologie più rigorose dal punto di vista matematico, idonee a prendere in considerazione anche ulteriori informazioni sui crimini quali: ora e giorno in cui essi si verificano; prossimità temporale rispetto ad un altro evento (ad esempio, un grande concerto); stagione; condizioni meteorologiche; intervallo di tempo fra gli episodi, per quanto riguarda i crimini seriali; ripetizione di un dato luogo; disposizione spaziale degli incidenti; progressione geografica di una serie di crimini; tipo di luogo in cui accadono i crimini (ad esempio, un parco); coordinate geografiche (anche non precise per latitudine e longitudine, ma del tipo "vicino alla fermata del bus n° x"); fattori ambientali e target (ad esempio, luminosità o condizioni di vita del vicinato); caratteristiche economiche e demografiche del luogo del crimine (Du et al., 2013).

Tutte queste caratteristiche, da sole o combinate, hanno un valore predittivo nel breve e nel lungo periodo, sia nell'analisi dei crimini di per sé che nell'analisi degli hot spot in generale. Gli analisti spesso fanno previsioni partendo da dette caratteristiche spazio-temporali, usando metodi di statistica descrittiva, inferenziale e bivariata. Nel tempo, sono stati sviluppati diversi metodi, schematicamente riassunte di seguito, per condurre le analisi spazio-temporali (Chen et al., 2005):

1. mediante l'utilizzo delle *heat maps* (Figura 16), ossia delle tavole che mostrano, tramite diverse intensità di colori, le frequenze relative dei crimini, a date, tempi e condizioni ambientali differenti. È possibile creare delle *heat maps* con l'ausilio di software largamente diffusi come Excel o di altri software specialistici;
2. attraverso l'uso dei *spatio-temporal generalized additive models* (ST-GAM) e dei *local spatio-temporal generalized additive models* (LST-GAM), sviluppati da Wang e Brown (2012). Questi modelli rappresentano estensioni dei modelli di regressione sviluppati sui reticolati geografici, in cui i dati in *input* includono le probabilità, associate ad ogni cella del reticolo, che queste abbiano una particolare struttura spazio-

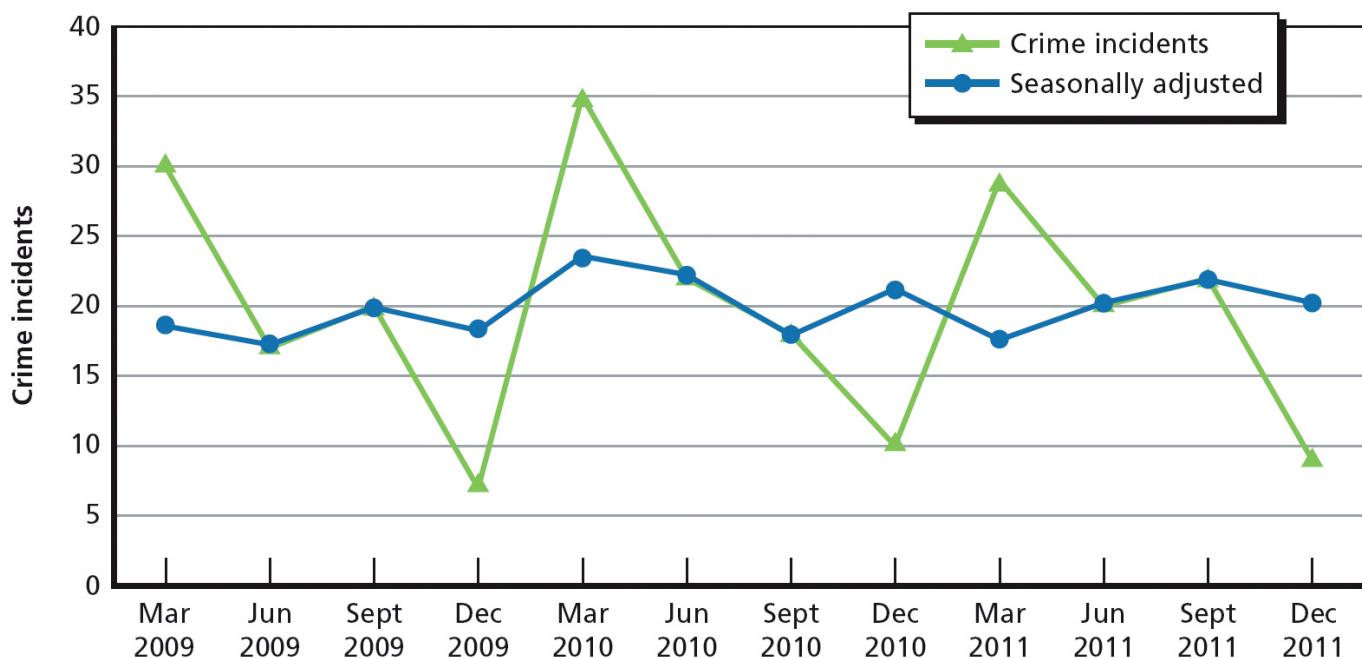
temporale in un dato tempo. In questo caso, una "caratteristica spazio-temporale" può essere un crimine precedente, un attributo generale o una determinata caratteristica geospaziale all'interno di una cella. Tutte queste caratteristiche possono essere indicizzate anche riguardo al tempo. Infatti, entrambi i modelli citati combinano le caratteristiche spazio-temporali dell'area dei crimini con i dati sul numero di crimini al fine di predire sia la *location* sia il tempo dei crimini futuri. Un ST-GAM è disegnato per predire i crimini per un'intera regione d'interesse. Questi modelli partono dall'assunzione che i *pattern* sottostanti il modello siano gli stessi per l'intera regione. I modelli LST-GAM, invece, lavorano per differenziati *patterns* regionali. Tutti e due i modelli calcolano la probabilità che il crimine analizzato sarà commesso in un determinato luogo ad un dato tempo, condizionatamente alle caratteristiche spazio-temporali dell'area in cui i crimini passati sono stati commessi. Wang e Brown (2012) hanno confrontato i risultati ottenuti con i due modelli con quelli ottenuti mediante un modello lineare generalizzato spaziale (GLM) e con quelli ricavati con la *hot spot analysis*. È risultato che i modelli ST-GAM e LST-GAM funzionano in maniera più efficiente rispetto ai GLM e ai metodi *hot spot*, ottenendo risultati più accurati.

**Figura 16. Heat maps** dei furti in appartamento e delle rapine nella città di Washington D.C

Hour	Six Part 1 Crimes							Burglaries							Robberies						
	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat
12:00 a.m.	829	483	464	421	480	546	874	45	40	37	38	37	38	57	133	93	86	56	83	88	169
1:00	634	284	298	278	297	321	677	41	26	32	32	32	27	34	132	66	78	55	59	82	162
2:00	534	223	213	202	231	243	535	30	27	34	28	33	28	34	144	42	50	39	55	45	141
3:00	477	168	140	133	141	173	472	26	29	26	28	22	31	31	159	21	34	25	27	49	155
4:00	334	163	156	155	141	169	317	33	31	31	30	28	43	36	86	27	26	30	29	24	89
5:00	136	100	104	107	94	112	115	13	17	19	16	15	19	14	29	23	17	19	15	22	26
6:00	140	240	241	208	220	210	214	18	50	47	35	47	49	28	23	25	27	17	29	21	32
7:00	275	612	602	574	612	600	342	35	179	172	156	187	147	44	22	28	28	24	19	27	35
8:00	183	390	358	388	402	393	192	10	105	91	91	88	96	16	23	22	15	27	29	17	21
9:00	313	609	556	540	524	547	380	35	109	119	111	113	113	45	12	37	31	30	19	31	25
10:00	413	486	483	424	462	480	434	39	78	83	91	72	79	44	29	36	40	25	51	27	28
11:00	424	485	463	473	475	498	483	35	88	61	69	81	79	53	35	44	33	43	46	55	36
12:00 p.m.	602	603	614	600	587	658	593	48	80	77	85	68	79	63	52	47	47	46	41	46	42
1:00	455	530	468	468	530	534	506	48	56	56	52	57	50	47	48	57	56	50	55	52	65
2:00	567	587	523	560	503	611	622	52	70	47	66	54	84	58	48	54	53	64	39	56	61
3:00	603	556	589	606	582	699	673	52	64	56	60	58	85	82	58	68	78	74	83	80	64
4:00	582	607	621	669	559	744	654	40	57	57	63	55	84	75	45	81	83	90	62	79	44
5:00	599	740	634	665	649	889	657	46	82	75	62	60	126	62	58	96	56	78	76	74	55
6:00	640	851	753	767	754	932	723	48	51	52	67	77	103	72	59	98	69	80	68	87	66
7:00	657	790	781	754	754	871	715	38	61	59	64	57	74	53	80	95	64	91	72	92	83
8:00	691	766	710	729	703	822	721	27	54	43	47	46	55	49	98	112	97	105	81	116	84
9:00	665	758	741	676	667	808	815	41	52	54	44	41	52	44	120	140	127	120	96	146	125
10:00	709	678	678	717	721	960	924	37	47	52	40	33	44	53	113	118	124	153	128	144	147
11:00	600	608	580	625	668	953	956	42	50	35	52	43	58	43	107	118	102	113	108	148	137

Fonte: *Information Builders, Law Enforcement Analytics Interface*, già pubblicato in RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013, p. 47

Figura 17. Spatio-temporal analysis. Trend dell'incidenza dei crimini



Fonte: RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013, p. 50

Nella spatio-temporal analysis assume rilevanza il fatto che alcuni tipi di crimini sono fortemente influenzati da *pattern* ciclici, come ad esempio il giorno della settimana o la stagione. Includere gli effetti ciclici in questo tipo di analisi è importante per eliminare le fonti conosciute delle variazioni, e così arrivare a migliori predizioni. Ai fini del calcolo della stagionalità, viene considerato adeguato un arco temporale della durata di cinque anni. Per gli effetti riguardanti il giorno della settimana, è accettabile considerare come arco temporale le cinque settimane precedenti. Produrre “trend aggiustati” sulla base degli effetti stagionali equivale a scomporre i dati sui crimini in *pattern* sistematici. Ad esempio, se si esaminano cinque anni di crimini su base quadriennale, si inizia con il calcolare la media del tasso di incidenza dei crimini per l'intero periodo dei cinque anni. Questo fornisce il *trend* dell'intero periodo, ossia il primo *pattern*. Poi, il secondo *pattern* deriva dal residuo, ossia la differenza tra il *trend* appena calcolato e l'incidenza dei crimini in ogni quadriennio. Il fattore stagionale è calcolato individuando la media quadriennale lungo i cinque anni di dati. Infine, il *pattern* stagionale aggiustato è calcolato sottraendo il fattore stagionale quadriennale dalla serie originale, per i corrispondenti quadrienni. In casi come quello appena prospettato, ci si riferisce generalmente come ad una *decomposizione additiva* e si assume che i dati siano indipendenti (Figura 17).

Altri metodi spazio-temporali disponibili sono inclusi nell'US Census Bureau's X-12-ARIMA, un modello che usa un software di aggiustamento stagionale con molte opzioni stagionali e di *trend*, e il TRAM/SEATS, un programma di aggiustamento stagionale sviluppato dalla Bank of Spain (Du et al., 2013).

Esistono diversi test per valutare la precisione dei modelli spazio-temporali (Johnson and Bowers 2004 a, b; Bowers and Johnson 2005; Johnson et al. 2007, 2009; Grubesic and Mack 2008). Questi, vengono presi in prestito dall'epidemiologia, ma vengono comunemente applicati anche in ambito criminologico. È il caso del test di Knox (Townsley et al., 2003) e del test di Mantel (Johnson and Bowers 2004 a, b). Il vantaggio di questi test è che possono quantificare il valore sulla scala spazio-tempo in cui avvengono le interazioni. Inoltre, i risultati di questi test possono essere visualizzati graficamente. D'altro canto, però, un punto a sfavore è dato dal fatto che con questi test non è semplice quantificare il cambiamento nel tempo, poiché gli *output* derivanti da essi forniscono una descrizione statica dei risultati.

L'analisi spazio-temporale mira ad analizzare l'evoluzione nello spazio e nel tempo dei crimini. Il metodo include informazioni relative alle caratteristiche ambientali e temporali dell'area di occorrenza dei crimini.

## Risk terrain analysis

La *risk terrain analysis*, sviluppata da Joel Caplan e dai ricercatori della Rutgers University del New Jersey (USA), comprende una famiglia di tecniche statistiche che mirano ad identificare le caratteristiche geografiche che contribuiscono al rischio che un crimine avvenga. Fanno predizioni circa il rischio del crimine basato su quanto sono vicine geograficamente le variabili di rischio.

Mediante la *risk terrain analysis* si mira dunque a fare previsioni relative al rischio che avvenga un crimine futuro. Dal punto di vista di un ufficiale di polizia, l'*output* di un modello *risk terrain* sarà qualitativamente lo stesso di quello di un modello *hot spot*, descritto in precedenza (vedi p. 20): entrambi evidenziano le aree che potrebbero essere soggette ad un alto tasso di crimini nel prossimo futuro. Dalla prospettiva di un ricercatore, tuttavia, i metodi citati sono fra loro molto diversi. I metodi della *hot spot analysis* sono fondamentalmente tecniche di *clustering*, che etichettano le aree dove i crimini si sono verificati. La modellizzazione della *risk terrain analysis* è, invece, un approccio classificatorio che caratterizza il rischio di una regione per determinate tipologie di crimini, basandosi sulle sue caratteristiche geografiche. Di seguito, vengono presentati due esempi di *risk terrain analysis*, uno semplice e facile da implementare e uno più complicato dal punto di vista matematico (Caplan e Kennedy, 2011).

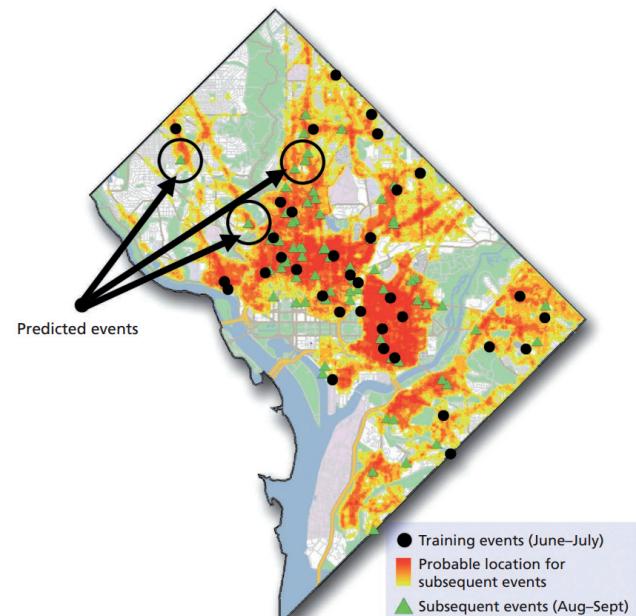
Il primo metodo è il *risk terrain modeling* (RTM), un approccio euristico, che viene utilizzato per valutare quanto i fattori geospaziali contribuiscono al rischio del crimine. Il RTM è implementabile attraverso l'uso di un pacchetto all'interno del software ArcGis. Il metodo consiste prima di tutto nella costruzione di un reticolo sulla zona geografica da analizzare. In seconda battuta, viene testata la relazione statistica tra la presenza di determinate caratteristiche geospaziali in ogni cella del reticolo e la presenza dei crimini di interesse all'interno delle celle. Le variabili che hanno una forte relazione positiva con il crimine vengono selezionate nel modello. Il metodo, poi, conta il numero di variabili selezionate presenti in ogni cella: le celle con il più alto numero di variabili che inducono i rischi vengono etichettate come *hot spot* (e tipicamente colorate in rosso o arancione). Il secondo metodo consiste in un approccio più propriamente statistico alla *risk terrain analysis* e si compone di due fasi. Nella prima fase, l'algoritmo compara le distanze fra i crimini e le caratteristiche geospaziali di interesse e traccia le distanze fra i crimini e le caratteristiche geospaziali più vicine, per ognuna delle tipologie di crimine e di caratteristica territoriale. Nella seconda fase, l'algoritmo valuta quanto "simile" sia ogni punto

sul reticolo rispetto alle *locations* che sono state teatro di crimini e rispetto alle distanze dalle caratteristiche geospaziali di interesse. I punti le cui distanze dalle caratteristiche geospaziali somigliano a quelli delle *location* dei crimini sono giudicati come ad alto rischio (Caplan e Kennedy, 2010).

La metodologia della *risk terrain analysis* ha due grandi vantaggi (Figura 18). Il primo è che questi metodi sono "genuinamente" predittivi, in quanto essi predicono il rischio sulla base degli attributi geografici, invece di estrararlo meramente dalla storia dei crimini pregressi. Matematicamente, ciò significa che questi metodi possono predire nuovi *hot spot* che sono simili anche in termini geospaziali agli altri *hot spot*. L'idea è che, poiché i nuovi *hot spot* non si basano sui dati relativi ai recenti crimini, sono abbastanza simili ai precedenti *hot spot* che sono stati considerati come ad alto rischio. Il secondo vantaggio è che tali strumenti possono mostrare quali tipi di caratteristiche geospaziali sono state usate nei modelli predittivi del crimine, al fine di condurre studi anche su quali tipologie di caratteristiche fisiche del territorio abbiano incidenza sul proliferare della criminalità.

Nell'uso delle metodologie *risk terrain*, si deve procedere con cautela, poiché, nel voler prevedere i nuovi *hot spot* correlati ai crimini recentemente avvenuti, si può facilmente incorrere nell'errore di generare mappe che mostrano come *hot* la zona più popolata di un'area geografica (Caplan e Kennedy, 2010; 2011).

**Figura 18.** Mappa sviluppata con la *Risk Terrain Analysis*, utilizzata per predire il rischio di borseggio nella città di Washington D.C. nell'estate 2008



Fonte: P. Borissow, *Crime Forecast of Washington D.C.*, Wikimedia Commons public domain image, 30 Marzo 2009

**La *risk terrain analysis* identifica le caratteristiche geografiche che contribuiscono al rischio che un crimine avvenga.**

## Agent based modelling

L'*agent based modelling* (ABM) rappresenta uno degli ultimi sviluppi in seno ai metodi statistici per la previsione della concentrazione della criminalità. Tale metodologia costituisce un'evoluzione del *cellular automaton* (CA) model, sviluppato a partire dagli anni '50 nell'ambito delle scienze computazionali (Von Neumann e Burks, 1966).

Mediante l'ABM, si punta a comprendere il comportamento degli "agenti", che possono essere sia individui sia gruppi di individui, e a scoprire le possibili interazioni tra essi all'interno di uno specifico territorio, al fine di valutarne gli effetti sul sistema nel suo complesso. In ambito criminologico, l'*agent based modelling* è utilizzato per fare previsioni circa il rischio di un crimine futuro, sulla base delle interazioni che avvengono in un determinato contesto tra diversi soggetti (nello specifico tra criminali e vittime). In particolare, esso attinge alle teorie razionali (principalmente alla teoria delle attività di routine) e alla criminologia ambientale<sup>7</sup>.

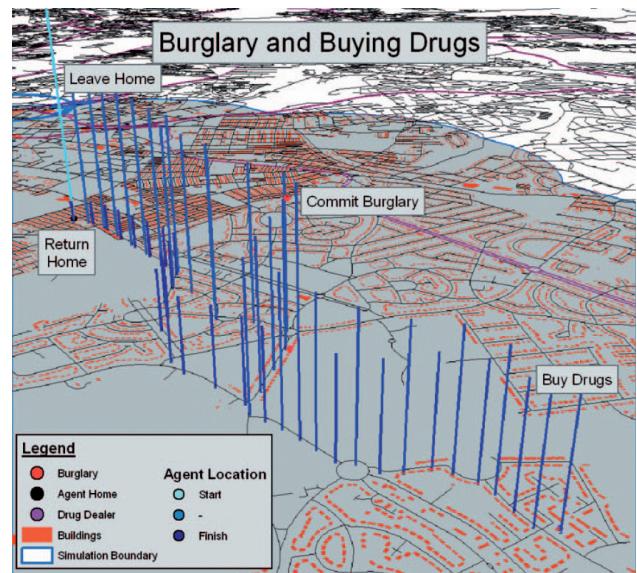
Un *agent based model*, nella sua versione classica, si compone di:

1. un agente, che viene caratterizzato attraverso una serie di variabili (come ad esempio, età, genere, status sociale) e di un "codice", che mira a rappresentare le azioni che tale agente porrà in essere (ad esempio, muoversi, fermarsi);
2. un modello, che specifica le modalità con cui si comporterà un determinato agente;
3. un ambiente, entro cui gli agenti agiscono ed interagiscono.

Uno dei pregi dell'*agent based modelling* è rappresentato dalla relativa semplicità di implementazione. Infatti, molti ABM, oltre ai suddetti elementi, ricorrono all'utilizzo di ulteriori modelli cognitivi che vanno ad integrare i processi comportamentali dell'agente. Un esempio in tal senso è rappresentato dal BDI (*Belief Desire Intention*) model (Wooldridge, 2009), in cui l'agente prende le decisioni sulla base dell'utile che ne può derivare.

Anche tale metodologia, tuttavia, non è esente da critiche: una delle principali riguarda la natura dell'ABM, ovvero se debba essere considerata astratta, essendo impossibile riprodurre tutte le variabili in gioco in un contesto con più agenti, oppure realistica, tenuto

Figura 19. Gli spostamenti criminali dell'agente



Fonte: N. Malleson et al., *Implementing comprehensive offender behaviour in a realistic agent-based model of burglary*, 2012

presente che alla base delle analisi vi è comunque l'utilizzo di dati "reali", come quelli di polizia o quelli provenienti dalle indagini di vittimizzazione. La risposta a tale problematica risiede nella quantità e nella qualità dei dati utilizzati per la strutturazione del modello: infatti, quanto più saranno vari e dettagliati, tanto più l'ABM potrà essere concreto.

Tale affermazione conduce ad un'altra critica che viene mossa nei confronti dell'*agent based modelling*, ossia che un modello molto dettagliato, per poter funzionare, avrà bisogno di un gran quantitativo di RAM e di tempo di calcolo. Nonostante tali difficoltà pratiche, l'ABM si è dimostrato un metodo attendibile per la previsione dei reati futuri, specie se in presenza di una grande disponibilità di dati e risorse.

Il sistema più utilizzato per costruire un modello *agent based* è NetLogo, per via della sua facilità di programmazione e della compatibilità con gli altri software comunemente utilizzati nell'ambito delle analisi statistiche, come R, Weka e Pajek (Malleson et al., 2012).

**L'*agent based modelling* si è dimostrato un metodo attendibile per la previsione dei reati futuri, specie se in presenza di una grande disponibilità di dati.**

<sup>7</sup> Perchè il crimine è prevedibile, p. 7.

## Conclusioni

Questa rassegna dei metodi predittivi ha inteso descrivere i principali modelli utilizzati nei recenti progetti di *predictive policing* e di *predictive urban security*. Il capitolo ha dapprima illustrato i metodi *hot spot*, ovvero quelli maggiormente diffusi ed utilizzati dai dipartimenti di polizia, che però mancano di sofisticatezza. Successivamente, sono stati descritti i metodi di *data mining*, che vengono usati quando le informazioni sono sufficienti ed è disponibile un software di *data mining*; questi metodi si rivelano molto interessanti per segnalare i *pattern* dei crimini. Di seguito, sono stati presentati i metodi di regressione, ossia quelli usati dai dipartimenti che hanno a disposizione una grande quantità di dati, così da produrre risultati maggiormente significativi dal punto di vista statistico. Inoltre, sono stati tratteggiati i metodi *near-repeat*, che si basano sulla proposizione che i crimini futuri avverranno vicino ai crimini attuali, e i metodi spazio-temporali, che, invece, mirano ad evidenziare graficamente le caratteristiche temporali e geografiche dei reati. La parte finale del rapporto è dedicata alla *risk terrain analysis*, con la quale si mira ad indagare quali siano i fattori che aumentano il rischio criminale, con una particolare attenzione posta alla vicinanza temporale con crimini passati e all'*agent based modelling*, che indaga le conseguenze delle interazioni nello spazio tra gli agenti, per capire dove è più probabile che avvenga un evento criminale.

Al termine di questo rapporto, che sintetizza i metodi statistici maggiormente utilizzati nei progetti di *predictive policing* e *predictive urban security*, è importante notare come la scelta della metodologia da utilizzare sia largamente influenzata dalla dimensione dei *data set* di cui dispongono le forze di polizia: di fronte ad una scarsità di dati, infatti, si utilizzano metodi più semplici, euristici, come quelli della *hot spot analysis*, in quanto richiedono strumenti poco complessi; quando vi è ampia disponibilità di informazioni, invece, si può fare ricorso a metodi analitici più sofisticati ed impegnativi, in termini di costi e di personale. È fondamentale precisare che tutte le metodologie presentate non predicono dove e quando il prossimo crimine sarà commesso, ma solo il livello di rischio criminale che sarà associato ad un particolare posto ad una determinata ora.

## Concetti fondamentali

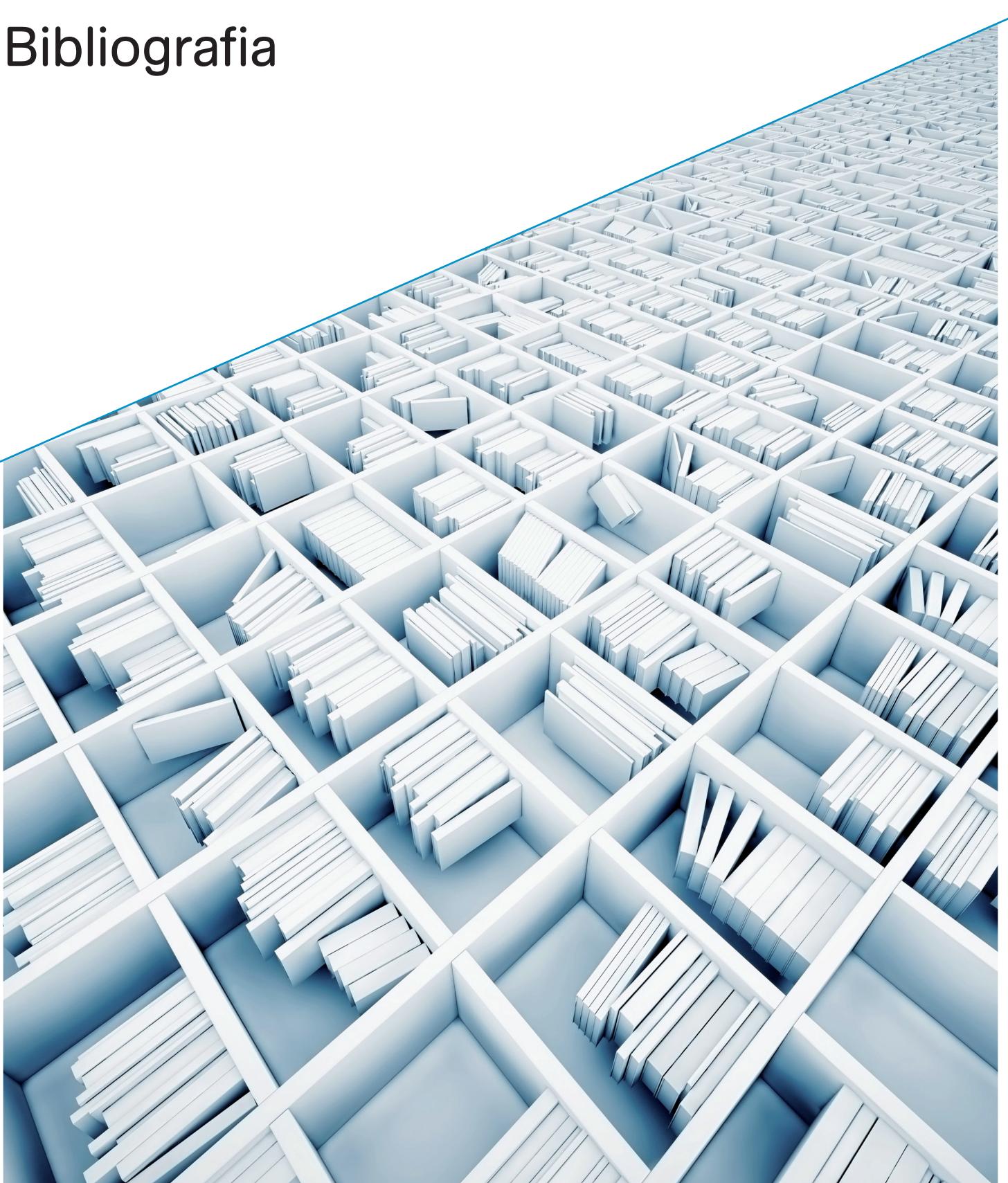
1. La capacità del *predictive policing* di mettere a disposizione informazioni qualitativamente migliori nell'ambito dell'attività di polizia può consentire alle forze dell'ordine di contrastare il crimine più efficacemente, giocando d'anticipo. Allo stato attuale, i progetti di polizia predittiva utilizzano differenti metodologie di analisi dei dati, a seconda dell'obiet-

tivo che si vuole raggiungere e del *background* di informazioni di cui si dispone.

2. Mediante la *hot spot analysis*, si mira a predire le aree nelle quali, sulla base dei dati storici dei crimini, la concretizzazione di un reato può essere più frequente. Rientrano in tale categoria: il *grid mapping*, le *covering ellipses*, il *KDE* e tutti i metodi euristici (l'identificazione manuale, il *quadrat thematic mapping*, le *jurisdiction-bounded areas*).
3. Mediante i metodi di *data mining*, si mira ad individuare all'interno di grandi set di dati dei *pattern* e dei *trend*. All'interno del *data mining*, si distinguono: i metodi di classificazione, il *clustering*, i metodi *ensamble*.
4. Mediante i *regression methods*, si punta a stimare il rischio che si verifichi un crimine in futuro mettendo in relazione la variabile da prevedere, ossia il pericolo che venga posto in essere un comportamento illecito, con le variabili esplicative, costituite non solo dai dati relativi ai crimini passati, ma anche da un ampio set di informazioni ulteriori. I metodi di regressione si distinguono sulla base della relazione che intendono stimare e possono consistere in: regressione lineare, regressione non lineare, regressione *spline*, regressione spaziale.
5. Mediante i metodi *near-repeat*, si operano previsioni partendo dall'assunto che i crimini futuri si ripeteranno in zone molto vicine a quelle in cui sono avvenuti i crimini passati e presenti.
6. Mediante la *spatio-temporal analysis* si vogliono analizzare le caratteristiche spaziali e temporali riguardo la *location* dei crimini. L'idea, infatti, è che queste informazioni siano estremamente utili per predire i futuri *hot spot*, anche mediante metodi analitici di misurazione.
7. Mediante la *risk terrain analysis* si mira ad identificare le caratteristiche geografiche che contribuiscono ad aumentare il rischio che avvenga un crimine in futuro.
8. Mediante l'*agent based modelling* si punta a comprendere quali interazioni tra gli agenti, in un contesto specifico, aumentino il rischio criminale.
9. La scelta della metodologia da utilizzare per predire le future concentrazioni dei crimini è influenzata dal *data set* di cui dispongono le forze dell'ordine o i gruppi di ricerca, che intendono porre in atto un progetto di *predictive policing*: infatti, con scarsa disponibilità di dati si utilizzano metodi più semplici; con grande disponibilità di dati, viceversa, si ricorre a metodi analitici più sofisticati ed impegnativi.

b

# Bibliografia



## Capitolo 1

- BECK C., McCUE C., *Predictive policing: what can we learn from WalMart and Amazon about fighting crime in a recession?*, in *The Police Chief*, novembre 2009. Reperibile al sito internet [www.policechiefmagazine.org/magazine](http://www.policechiefmagazine.org/magazine) (data ultima consultazione 02/10/2014).
- BRAGA A.A., *Problem-Oriented-Policing and crime prevention*, New York, 2002.
- BRANTINGHAM P.J., BRANTINGHAM P.L., *Environmental criminology*, Chicago, 1991.
- BRANTINGHAM, P.J., BRANTINGHAM, P.L., "Criminality of Place: Crime Generators and Crime Attractors", *European Journal on Criminal Policy and Research*, vol. 3, n. 3, 1995.
- CHIESI, L., "L'ipotesi delle inciviltà. La non ovvia relazione tra manutenzione e senso di insicurezza", in Amendola G. (a cura di), *Il governo della città sicura. Politiche, esperienze e luoghi comuni*, Liguori, Napoli, 2003.
- CLARKE R.V., "Introduction", in Clarke R.V. (a cura di), *Situational Crime Prevention. Successfull Case Studies*, New York, 1997.
- CORNISH D., CLARKE R., *The reasoning criminal: rational choice perspectives on offending*, New York, 1986.
- DI NICOLA A., BRESSAN S., "Sicurezza urbana predittiva: eSecurity e le nuove prospettive per la prevenzione della criminalità nelle città", in *Sentieri Urbani*, 13, 2014.
- DULLI F., FURINI F., PERON E., *Data mining: metodi e strategie*, Milano, 2009.
- FELSON M., CLARKE R.V., *Opportunity makes the thief. Practical theory for crime prevention*, London, 1998.
- FELSON M., COHEN L., *Social change and crime rate trends: A routine activity approach*, in *American sociological review*, Chicago, 1979.
- GOLDSTEIN H., *Improving policing: a Problem-Oriented approach*, New York, 1979.

- GOLDSTEIN H., *What is POP?*, [www.popcenter.org/](http://www.popcenter.org/) (data ultima consultazione 02/10/2014).
- HARRIES K., *Mapping crime. Principles and practice*, Washington, 1999.
- LAB S.P., *Crime Prevention: Approaches, Practices, and Evaluations – 7th Edition*, New Providence, Matthew Bender & Company, 2010.
- LYON D., *La società sorvegliata. Tecnologie di controllo della vita quotidiana*, Milano, 2003.
- NAISBITT J., *Megatrends: ten new directions transforming our lives*, New York, 1982.
- NOBILI G., "Disordine urbano e insicurezza: una prima indagine su Bologna", in *Quaderni di Città Sicure – Regione Emilia Romagna*, n. 28, 2003.
- RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013.
- RATCLIFFE J.H., *Intelligence-Led Policing*, Collumpton, 2008.
- REGIONE PIEMONTE, *Leggere la sicurezza. I dati, il contesto, i fenomeni e le percezioni*, Regione Piemonte, Torino, 2012.
- ROUFA T., *Learning to Predict When and Where Crime Will Occur*, [www.criminologycareers.about.com](http://www.criminologycareers.about.com) (data ultima consultazione 02/10/2014).
- SARTORI L., *La società dell'informazione*, Bologna, 2012.
- SCHONBERGER V., CUKIER K., *Big Data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, Milano, 2013.
- SHERMAN L., *Hot spot of crime and criminal careers of places*, in *Crime and place. Crime prevention studies Vol.1*, a cura di J. Eck – D. Weisburd, New York, 1995.

VEZZADINI S., *Profilo geografico e crime mapping. Il contributo della criminologia ambientale allo studio del delitto*, in *Scena del crimine e profili investigativi: quale tutela per le vittime?*, a cura di R. Bisi, Milano, 2006.

UCHIDA C., *A national discussion on predictive policing: defining our terms and mapping successful implementation strategies*, [www.ncjrs.gov](http://www.ncjrs.gov) (data ultima consultazione 02/10/2014).

WARTELL J., GALLAGHER K., "Translating environmental criminology theory into crime analysis practice", *Policing. A Journal of Policy and Practice*, vol. 6, n. 4, 2012.

WILLIAMS F.P., Mc SHANE M.D., *Devianza e criminalità*, Bologna, 2002.

WILSON J.Q., KELLING G.L., "Broken Windows", in *The Atlantic Monthly*, vol. 279, n. 3, 1982.

## Capitolo 2

ARMSTRONG T., *Managing for 21st Century Crime Prevention in Memphis*, [www.managementexchange.com](http://www.managementexchange.com) (data ultima consultazione 03/10/2014).

BOWERS J., JOHNSON D., *Domestic Burglary Repeats and Space-Time Clusters. The Dimensions of Risk*, in *European Journal of Criminology*, Sheffield, 2005.

CHAINY S., *Predictive mapping (predictive policing)*, 2012, [www.ucl.ac.uk](http://www.ucl.ac.uk) (data ultima consultazione 03/10/2014).

FRIEND Z., *Predictive Policing: Using Technology to Reduce Crime*, in *FBI Law Enforcement Bulletin*, settembre 2013, [www.fbi.gov](http://www.fbi.gov) (data ultima consultazione 03/10/2014).

GARRETT R., *Predict and serve*, [www.officer.com](http://www.officer.com) (data ultima consultazione 03/10/2014).

JONES V., *Disrupting the optimal forager: predictive risk mapping and domestic burglary reduction in Trafford, Greater Manchester*, [www.ucl.ac.uk](http://www.ucl.ac.uk) (data ultima consultazione 03/10/2014).

MULLER J.K., *Preview of Predictive Policing. IBM Spatial Solutions*, [https://www.ibm.com/developerworks/community/blogs/johnmuller/entry/preview\\_of\\_predictive\\_policing\\_publication11?lang=en](https://www.ibm.com/developerworks/community/blogs/johnmuller/entry/preview_of_predictive_policing_publication11?lang=en) (data ultima consultazione 03/10/2014).

VLAHOS J., *Come anticipare il crimine*, in *Le Scienze*, Roma, febbraio 2012.

## Capitolo 3

ANSELIN L., HUDA S., *Spatial econometrics in practice: a review of software options*, in *Regional Science and Urban Economics*, 1992.

BALLER R., *Structural covariates of US county homicide rates: incorporating spatial effects*, in *Criminology*, 2001.

BERNASCO W., ELFFERS H., *Statistical analysis of spatial crime data*, in *Handbook of quantitative criminology*, Berlin, 2010.

BOWERS K., JOHNSON S., *Domestic burglary repeats and space-time clusters*, in *European Journal of Criminology*, 2005.

BUSHMAN B., "Is the Curve Relating Temperature to Aggression Linear or Curvilinear? Assaults and Temperature", in *Minneapolis Reexamined*, in *Journal of Personality and Social Psychology*, 2005.

CAPLAN, J.M., KENNEDY, L.W. (a cura di), *Risk Terrain Modeling Manual: Theoretical Framework and Technical Steps of Spatial Risk Assessment for Crime Analysis*, Newark, Rutgers Center on Public Security, 2010.

CAPLAN, J.M., KENNEDY, L.W. (a cura di), *Risk Terrain Modeling Compendium: For Crime Analysis*, Newark, Rutgers Center on Public Security, 2011.

CHAINY S., TOMPSON, L., "Engagement, Empowerment and Transparency: Publishing Crime Statistics using Online Crime Mapping", *Policing*, pubblicato online il 14 marzo 2012, (data ultima consultazione 02/10/2014).

CHAMONT, W., PIN-SHUO, L., "Data Mining and Hotspot Detection in an Urban Development Project", *Journal of Data Science*, 6: 389-414, 2008.

CHEN, H., CHUNG, W., XU, J.J., WANG, G., QIN, Y., CHAU, M., "Crime data mining: a general framework and some examples", *Computer*, 37/4: 50-56, 2004.

CHEN, D., WEEKS, J.R., KAISER, V.J., "Remote Sensing and Spatial Statistics as Tool in Crime Analysis", in Wang, F. (a cura di), *Geographic Information Systems and Crime Analysis*, Idea Group Publishing, 270-291, 2005.

CHUNG-HSIEN, Y., WARD, M.W., MORABITO, M., DING, W., "Crime Forecasting Using Data Mining Techniques", in *Data Mining Workshops (ICDMW)*, 2011 IEEE 11th International Conference. 11/11: 779-786, 2011.

- DE BRUIN, J.S., COCX, T.K., KOSTERS, W.A., LAROS, J.F.J., KOK, J.N., "Data Mining Approaches to Criminal Career Analysis", *Proceedings of the ICDM '06. Sixth International Conference*, 171–177, 2006.
- DU, P., XIA, J., DU, Q., LUO, Y., TAN, K., "Evaluation of the spatio-temporal pattern of urban ecological security using remote sensing and GIS", *International Journal of Remote Sensing*, 34: 848-863, 2013.
- ECK, J.E., CHAINY, S., CAMERON, J.G., LEITNER, M., WILSON, R.E., *Mapping Crimes: Understanding Hotspots*, Washington, National Institute of Justice, 2005.
- GRUBESIC T., MACK E., "Spatio-temporal interaction of urban crime", in *Journal of Quantitative Criminology*, 2008.
- JIANQIANG, S., BOCHMANN, G., ADAMS, C., *A trust model with statistical foundation*, New York, Springer, 2005.
- JOHNSON S., BOWERS K., "The burglary as clue to the future", in *European Journal of Criminology*, 2004 (a).
- JOHNSON S., BOWERS K., "The stability of space-time clusters of burglary", in *British Journal of Criminology*, 2004 (b).
- JOHNSON S., BERNASCO W., BOWERS K.J., ELFFERS H., RATCLIFFE J., RENGERT G., TOWNSLEY M., "Space-time patterns of risk: a cross national assessment of residential burglary victimization", in *Journal of Quantitative Criminology*, 2007.
- LEVINE N., *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations*, 2004.
- MALLESON N., SEE L., EVANS A., HEPPENSTALL A., *Implementing comprehensive offender behaviour in a realistic agent-based model of burglary*, 2012.
- MESSNER S., ANSELIN L., *Spatial analyses of homicide with areal data*, in *Spatially integrated social science*, New York, 2004.
- MOHLER G.O., SHORT M. B., BRANTINGHAM P. J., SCHÖENBERG F. P., TITA G.E., "Self-Exciting Point Process Modeling of Crime", in *Journal of the American Statistical Association*, 2011.
- PORTER J., PURSER C., *Social disorganization, marriage, and reported crime: a spatial econometrics examination of family formation and criminal offending*, in *Journal of Criminal Justice*, 2010.
- RAND, *Predictive policing. The role in crime forecasting in law enforcement operations*, Washington, 2013.
- SHERMAN L., *Hot spots of crime and criminal careers of places*, in *Crime and place. Crime prevention studies Vol.1*, a cura di J. Eck – D. Weisburd, New York, 1995.
- SHORT, M.B, D'ORSOGNA, R., PASOUR, V.B., TITA, G.E., BRANTINGHAM, P.J., BERTOZZI, A.L., CHAYES, L.B., "A statistical model of criminal behavior", *Mathematical Models Methods in Applied Sciences*, 18: 1249, 2008.
- SHORT, M.B, D'ORSOGNA, R., BRANTINGHAM, P.J., TITA, G.E., "Measuring and Modeling Repeat and Near-Repeat Burglary Effects", *Journal of Quantitative Criminology*, 25/3: 325-339, 2009.
- SHYAM, V.N., "Crime Pattern Detection Using Data Mining", *Proceedings of the Web Intelligence and Intelligent Agent Technology Workshops*, IEEE/WIC/ACM International Conference, 41-44, 2006.
- SMITH, S.C., BRUCE, C.V., *CrimeStat III User Workbook*, Washington: National Institute of Justice, 2010.
- SULLIVAN T.J., PERRY W.L., "Identifying Indicators of Chemical, Biological, Radiological, and Nuclear (CBRN) Weapons Development Activity in Sub-National Terrorist Groups", in *Journal of the Operational Research Society*, 2004.
- TOWNSLEY M., HOMEL, R., CHASELING, J., "Repeat Burglary Victimization: Spatial and Temporal Patterns", in *Australian and New Zealand Journal of Criminology*, 2000.
- TOWNSLEY M., HOMEL, R., CHASELING, J., "Infectious burglaries. A test of the near repeat hypothesis", in *British Journal of Criminology*, 2003.
- VON NEUMANN, J., BURKS A. W., *Theory of self-reproducing automata*, Urbana, University of Illinois Press, 1966.
- WANG X., BROWN D.E., "The Spatio-Temporal Modeling for Criminal Incidents", in *Security Informatics*, 2012.
- WOOLDRIDGE, M., *An Introduction to MultiAgent Systems*, 2009



Digital print: [www.rotooffset.it](http://www.rotooffset.it) - Trento, Italy

Le opinioni espresse nel presente rapporto di ricerca sono di responsabilità esclusiva degli autori e non riflettono necessariamente la posizione ufficiale dell'Unione europea.

Trento, Dicembre 2014

© 2014 eCrime - Università degli Studi di Trento





