

Sicurezza | Modelli | Hotspot

# Algoritmi in campo per prevenire il crimine

Possibile sfruttare l'analisi dei Big data. Oppure ci si concentra sulle cause per determinare dove e quando il rischio è maggiore  
di **Andrea Carobene**

✪ Gli abitanti di Modesto, città della California con 210.000 abitanti, dormono sonni più tranquilli. Qui, lo scorso anno, le rapine e i furti di auto sono calati del 14% e quelli nelle abitazioni dell'11 per cento. Modesto è una tra le tante città che sfruttano le potenzialità dell'informatica per rendere più efficiente il lavoro della polizia. Tra queste località vi sono Los Angeles, Atlanta, Chicago e la britannica Medway. L'idea è quella di fare in modo che gli agenti siano presenti dove si sta per compiere un crimine, riuscendo così a prevenirlo. Il "trucco" è quello di usare dei software che, sulla base di dati statistici, prevedono i reati.

Uno di questi programmi è PredPol, realizzato da un gruppo di ricercatori delle Università di Santa Clara e della California di Los Angeles. PredPol viene descritto da Heather Graves, portavoce del dipartimento della Polizia di Modesto, come «strumento valido che usiamo per combattere il crimine». Il software analizza 10 anni di statistiche sui reati fornendo come risultato «aree quadrate di 500 per 500 piedi (circa 152 metri) dove è più probabile che possa accadere un reato». Quelle che il software indica sono gli hotspot, le "zone calde", sulle quali concentrare il lavoro preventivo della polizia. Come spiega Graves, il programma è usato per indirizzare «gli agenti nelle aree dove occorre essere presenti per ridurre il crimine». La presenza dei poliziotti «ha un effetto deterrente su reati come rapine e furti di appartamento», e in molti casi favorisce gli arresti.

Lo scorso 15 gennaio, sul *Journal of the American Statistical Association*, Jeffrey Brantingham con altri colleghi ha spiegato nel dettaglio il funzionamento del programma. Brantingham insegna all'Università della California, è tra gli ideatori di Pre-

dPol, ed è colui che ne ha seguito l'implementazione per la polizia di Los Angeles, come conferma il capitano Ernest Eskridg, capo del distretto Foothill della città californiana. L'algoritmo alla base del funzionamento di PredPol si fonda sulla distribuzione di Poisson, funzione statistica descritta più nel dettaglio nell'articolo a fianco. I ricercatori hanno verificato l'attendibilità del loro algoritmo analizzando l'attività delle polizie di Los Angeles e del Kent, in Inghilterra. Seguendo le indicazioni del programma, l'efficacia delle pattuglie cresce anche di due volte e si ottiene una riduzione dei reati del 7,4 per cento.

Sono diversi i programmi che utilizzano metodologie simili. Tra questi CrimeStat: un software statistico di analisi del crimine che può essere scaricato gratuitamente dal sito del ministero della Giustizia statunitense. Altri programmi di questo tipo sono Blue Crush di Ibm o Risk Terrain Modeling, sviluppato all'Università Rutgers nel New Jersey. Come spiega Marco Dugato, che insegna Metodi e tecniche della ricerca criminologica all'Università Cattolica di Milano, occorre tuttavia distinguere fra due diversi approcci. In un caso «si lavora sui crimini avvenuti nel passato e, attraverso un algoritmo, si studia la probabilità che uno specifico crimine possa avvenire in una determinata area dopo un certo tempo». È quanto avviene con un programma come PredPol. Nel secondo approccio, invece, «ci si concentra sulle cause del fenomeno per determinare dove e quando il rischio è maggiore».

Lo scorso anno il Centro Transcrime dell'Università Cattolica di Milano, diretto dal professor Ernesto Savona, pubblicò per il ministero dell'Interno un lavoro sulla previsione dei furti nelle abitazioni. Lo studio, realizzato da quattro ricercatori tra cui Dugato, partiva proprio dalle cause: «Noi sappiamo che i reati si concentrano, e che questi hotspot non sono completamente casuali - chiarisce Dugato - ma variano sulla base di alcuni fattori come il contesto urbanistico o le condizioni socioeconomiche di una determinata area. Dopo aver identificato i fattori di rischio e quelli protettivi, il

modello di Transcrime seleziona quelli più significativi per costruire le mappe previsionali. Lo studio è stato realizzato su tre città: Milano, Roma e Bari, sfruttando informazioni accessibili con sistemi di open data come le distribuzioni demografiche, le statistiche sul reddito, ma anche la presenza di case popolari o i valori immobiliari. Per verificare l'attendibilità delle previsioni i ricercatori hanno finto di essere nel 2013 e di prevedere cosa sarebbe avvenuto nel 2014, per poi paragonare i loro risultati con le statistiche sui furti realmente avvenuti. Il risultato è stato interessante: a Milano con questo metodo sarebbe stato possibile prevedere il 29,3% dei furti, a Bari il 41,6% e a Roma addirittura il 47,1%, quasi uno su due. Esperimenti di questo tipo dimostrano che con i "big data", è possibile migliorare l'efficacia nella lotta al crimine, come è già avvenuto ad esempio nella Questura di Milano dove il software KeyCrime più volte si è rivelato efficace nella prevenzione di alcuni reati.

«Questi programmi - continua Dugato - possono essere usati per analizzare altri fenomeni criminali». Ad esempio, il centro Transcrime «ha sviluppato una serie di modelli per prevedere il rischio di infiltrazioni della criminalità organizzata o del riciclaggio di denaro sporco, mentre a breve partirà un progetto per analizzare i fattori che favoriscono il reclutamento dei terroristi».

E proprio Milano è stato il capofila, due anni fa, di SmartCiber, un programma europeo che aveva come scopo quello di studiare il rischio terrorismo. Il progetto coinvolgeva più città europee e diverse società come A2A, Amsa, Atm o Sea, e si basava sulla mappatura georeferenziata delle infrastrutture critiche del territorio e l'individuazione dei fattori di rischio, tanto quelli relativi ad episodi specifici quanto quelli generati dalla presenza di particolari variabili socioeconomiche. «L'obiettivo - conclude Dugato - è quello di realizzare modelli che abbiano ricadute concrete», capaci di migliorare la sicurezza nelle città. Un obiettivo fondamentale, che la tecnologia può aiutare a raggiungere.

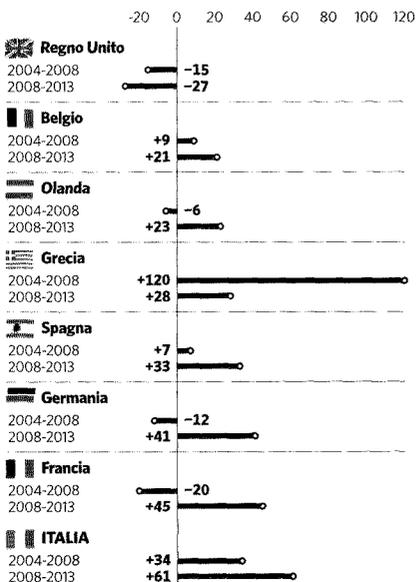
© RIPRODUZIONE RISERVATA

## I MODELLI MATEMATICI ANTI-CRIMINE

Sono due gli approcci. Uno è di tipo statistico, come il programma PrePol. In questo caso si lavora sui crimini avvenuti nel passato e attraverso un algoritmo si studia la probabilità che uno specifico crimine possa avvenire in una determinata area dopo un certo tempo. Il secondo, come il modello forecasting, si concentra sulle cause del fenomeno per determinare dove e quando il rischio è maggiore per costruire mappe previsionali

### TASSO DI FURTI IN ABITAZIONE

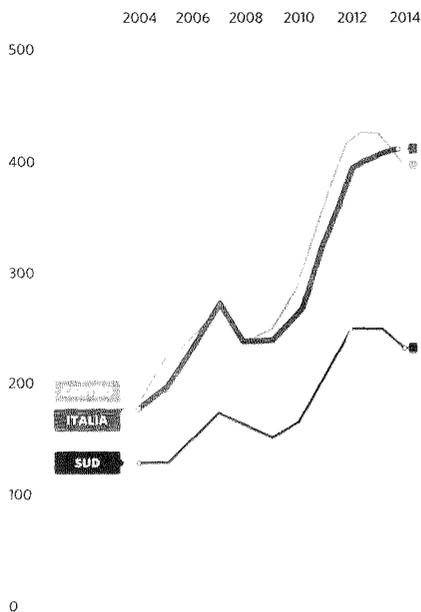
Var. % in otto Paesi europei, 2004-2013\*



\*È stato considerato come ultimo anno il 2013 perché il dato ufficiale 2014 non è al momento disponibile. Sono stati considerati gli 8 paesi europei con il tasso di furti in abitazione più alto all'inizio della serie storica

### LE CASE (SVALIGIATE) IN ITALIA

Andamento del tasso di furti in abitazione ogni 100 mila abitanti in Italia per macroarea, 2004-2014



Fonte: elaborazione Transcrime di dati SDI- Ministero dell'interno

### IL MODELLO FORECASTING

Fase 1

#### IDENTIFICARE I FATTORI DI RISCHIO E/O PROTETTIVI

Si sono individuati alcuni fattori che potrebbero facilitare o contrastare i furti in abitazione

#### FATTORI DI RISCHIO

##### Forte influenza

Alta densità abitativa  
 Alta densità di furti in abitazione nell'anno precedente  
 Alti valori immobiliari

##### Media influenza

Alta % della popolazione + 70 anni  
 Presenza di negozi "Compro oro"  
 Presenza di edilizia residenziale pubblica

##### Dobole - nulla influenza

Alta % di stranieri residenti  
 Bassa densità abitativa  
 Bassa percentuale di votanti  
 Bassi valori immobiliari  
 Presenza di campi nomadi  
 Presenza di sedi servizi per le tossicodipendenze

#### FATTORI PROTETTIVI

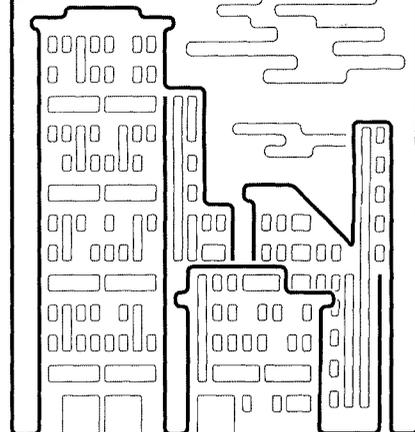
##### Debole - nulla influenza

Alta densità abitativa  
 Alta % popolazione + 70 anni  
 Alti valori immobiliari  
 Presenza di presidi Forze di Polizia

Fase 2

#### SELAZIONARE I FATTORI INFLUENTI

Sono stati selezionati solo i fattori che hanno dimostrato, tramite un algoritmo, di avere un potere predittivo



Fase 3

#### COSTRUIRE LE MAPPE PREVISIONALI

Ciascuna città è stata divisa in piccole aree della misura di 2.500 mq (l'equivalente di circa la metà di un campo di calcio)

A ogni area della città è stato associato un punteggio di rischio sulla base dei fattori selezionati nella fase 2

In questo modo è stato possibile identificare le aree più a rischio di ciascuna città

Fase 4

#### VALUTARE LA CAPACITÀ PREDITTIVA

Le mappe previsionali sono state create utilizzando i dati 2012 e 2013

Il risultato della previsione è stato confrontato con la distribuzione dei reati effettivamente denunciati nel 2014

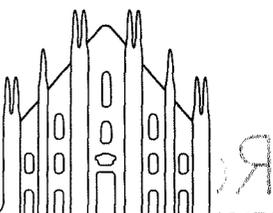
La corrispondenza tra la previsione e i reati realmente avvenuti ha permesso di valutare la capacità predittiva delle mappe previsionali

Le mappe previsionali specifiche per giorno e fascia oraria hanno predetto tra il 17,7% e il 44,6% dei furti in abitazione

Per ciascuna città sono state create anche 4 mappe previsionali secondo il giorno della settimana e l'orario

Sia i furti in abitazione che i fattori di contesto in grado di spiegarli potrebbero, infatti, cambiare nel tempo. È, quindi, necessario che le previsioni siano specifiche

# Milano



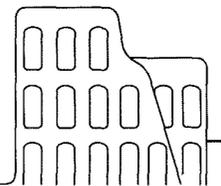
■ Rischio molto alto (2,2%) ■ Rischio alto (5,5%)



Le aree definite a rischio sono concentrate tra la circonvallazione interna e quella esterna della città e coinvolgono soprattutto i quartieri di Centrale, Buenos Aires - Porta Venezia, XXII Marzo, Isola e De Angeli-Monte Rosa



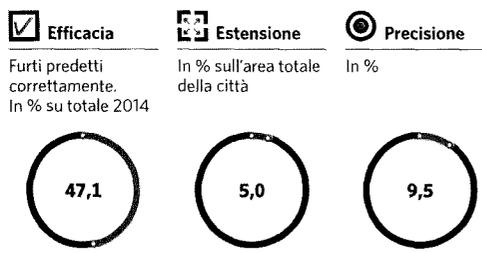
# Roma



■ Rischio molto alto (2,5%) ■ Rischio alto (2,5%)

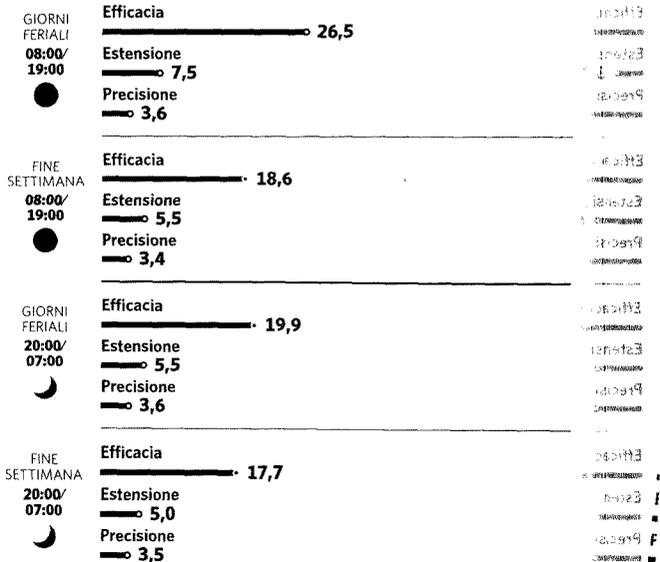


Le più a rischio si concentrano nel centro storico (rioni Ponte, Regola, Sant'Eustacchio, Sant'Angelo, Campo Marzio, Colonna ed Esquilino) e nei quartieri Aurelio, Gianicolense, della Vittoria, Appio Latino e Nomentano

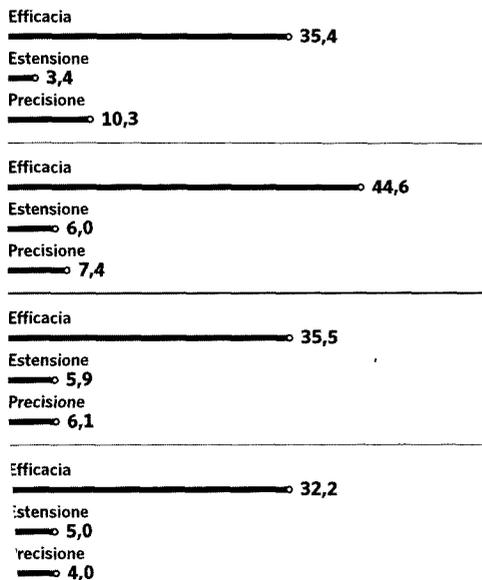


LO SCHEMA SEGUENTE RIASSUME LE 12 MAPPE CREATE E SINTETIZZA I RISULTATI OTTENUTI

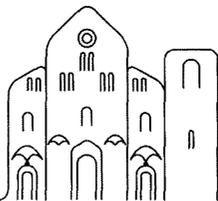
10 20 30 40 50 60



10 20 30 40 50 60



# Bari



■ Rischio molto alto (2,6%) ■ Rischio alto (1,5%)



Queste aree si concentrano principalmente nei quartieri Libertà, Murat e Madonnella

**Efficacia**  **Estensione**  **Precisione**  
 Furti predetti correttamente. In % su totale 2014



10 20 30 40 50 60

**Efficacia** 50,4

**Estensione** 5,0

**Precisione** 10,1

**Efficacia** 47,5

**Estensione** 4,7

**Precisione** 10,1

**Efficacia** 52,6

**Estensione** 5,4

**Precisione** 9,7

**Efficacia** 37,6

**Estensione** 5,1

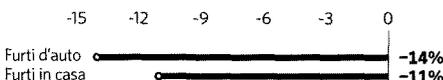
**Precisione** 7,3

## IL MODELLO STATISTICO

Le zone calde



La mappa è una griglia di aree quadrate (di circa 152 metri) che identificano le più alte concentrazioni di reati. Modesto è una delle tante città americane, come Los Angeles, Atlanta e Chicago, che sfruttano le potenzialità dell'informatica per rendere più efficiente il lavoro della polizia. L'anno scorso grazie al software che analizza 10 anni di statistiche sui reati, le rapine e i furti d'auto e quelli nelle abitazioni sono diminuiti

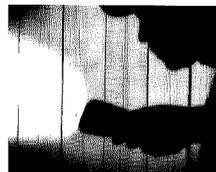


### A Santa Cruz in tempo reale

Mappe di questo tipo possono essere arricchite da qualunque informazione si consideri utile per arricchire i database:  
 informazioni anagrafiche  
 informazioni reddituali  
 illuminazione  
 tipologia di abitazione  
 dati raccolti dai cellulari o dallo spostamento dei veicoli

L'EFFICACIA DELLE PATTUGLIE RADDOPPIA  
 Riduzione dei reati **7,4%**

Legge di Poisson | Incidenti | Distribuzione



# Probabilità certe

La legge degli "eventi rari" può essere applicata a una miriade di fenomeni

● I cavalli possono essere pericolosi. Tra il 1875 e il 1894 questi animali uccisero infatti ben 196 soldati prussiani appartenenti a 14 reggimenti. Incuriosito da tali numeri, alla fine dell'800 il matematico russo Ladislaus Bortkiewicz si mise a studiare da un punto di vista statistico questi decessi. Bortkiewicz, che insegnava in Germania, costruì una tabella nella quale segnò tutte le 280 osservazioni eseguite, considerando come "osservazione" la lettura dei diari di un anno di un singolo reggimento. I diari che non riportavano alcun incidente erano 144; quelli che riportavano una vittima erano 91; e così via, fino a quelli con quattro vittime che erano due. Il matematico si accorse che questi numeri non erano casuali, ma seguivano con una precisione quasi stupefacente la legge di Poisson: una distribuzione che dà la probabilità di ottenere in un intervallo di tempo un certo numero di

eventi uguali quando si conosce la probabilità di uno solo di tali eventi.

La legge di Poisson può essere applicata a una miriade di fenomeni: dal numero di chiamate in arrivo a un centralino alla quantità di incidenti su una strada trafficata. In uno studio del 1946 la si usò ad esempio per calcolare la distribuzione con la quale le bombe tedesche erano cadute a Londra all'interno di aree quadrate di mezzo chilometro di lato.

È sufficiente quindi conoscere la probabilità generale di un "fatto" per descrivere in anticipo come gli eventi con quella probabilità si distribuiranno nel tempo. Questa legge, chiamata anche degli "eventi rari" è l'ideale per prevedere i crimini. Basta infatti conoscere la probabilità di un furto all'interno di un'area quadrata ogni dodici mesi, per poter dire quante sono le zone della città nelle quali ci si aspetta che in un anno vi siano due furti, tre o nessuno.

Per migliorare le previsioni si usano algoritmi più complessi che, pur partendo sempre da questa distribuzione, tengono

ad esempio conto delle modifiche causate da nuovi eventi. Il team di Erik Lewis, dell'Università della California di Los Angeles, applicando questo ultimo tipo di modello ha descritto con maggiore fedeltà la dinamica dei decessi violenti di civili in Iraq tra il 2003 e il 2007.

In Nigeria, invece, il gruppo di Ayo Emmanuel Olajuyigbe, dell'Università di Akure, ha analizzato la distribuzione del crimine in quella città di quasi 600.000 abitanti. In questo caso gli studiosi si sono serviti del software CrimeStat, anch'esso basato sulla distribuzione di Poisson. I ricercatori hanno spiegato sull'*International Journal of Criminology and Sociological Theory* che lo studio ha dimostrato la necessità di sviluppare un database dei crimini, e di formare i poliziotti all'informatica per migliorarne le capacità investigative. La conclusione è che occorre aprire la strada "all'utilizzo delle moderne tecnologie nella lotta al crimine" in tutta la Nigeria.

Un auspicio che parte dal lontano, dall'osservazione dei cavalli dell'esercito del re di Prussia. (an.car.)

Diagrammi | Strategie | Simulazioni

# Matematica & pizzini

La mafia si combatte anche con l'analisi dei network per ricostruire le relazioni

● La mafia si combatte anche con l'analisi dei network. Lo studio scientifico delle relazioni esistenti all'interno di un'organizzazione criminale permette infatti di svelare il ruolo svolto dai singoli delinquenti e capire quali sono le azioni di polizia più efficaci. Giacomo Fiumara, ricercatore di Informatica all'Università degli Studi di Messina, assieme ad altri colleghi ha analizzato dieci anni di informazioni sulle mafie della Sicilia settentrionale. Sono stati così studiati i tabulati telefonici "depurati di ogni informazione personale" per estrarre il network dei collegamenti e ricostruire il diagramma delle relazioni.

Contemporaneamente, i ricercatori hanno disegnato un altro network nel quale due persone si consideravano connesse quando «erano presenti nello stesso mandato di cattura, erano state viste insieme al

bar e così via», come spiega Fiumara.

Gli studiosi hanno scoperto che i due network erano differenti; ad esempio, tutte le persone del secondo grafo erano presenti nel primo, tranne sei. Il punto è che «i capi non comunicano col telefono» ma, come la cronaca ci insegna, utilizzano altri metodi meno intercettabili come i "pizzini": foglietti di carta che «non sono folclore ma costituiscono una misura di sicurezza».

Il lavoro dell'Università di Messina ha permesso di rivelare i nodi essenziali del network, ma ha anche cercato di capire le strategie migliori per combattere questo tipo di organizzazioni. «L'idea di fondo è stata quella di provare a simulare come reagisce un network di questo tipo di fronte a due tipologie di azione delle Forze dell'ordine», vale a dire la retata in grande stile o la cattura mirata di alcune persone.

Il risultato è stato che l'approccio seriale, ossia la cattura di un vertice dopo l'altro dell'organizzazione, sembra infliggere maggiori danni all'organizzazione criminale. «Si tratta solamente di una simulazio-

ne», precisa Fiumara, ma una simulazione che può offrire strategie utili per l'azione delle Forze dell'ordine.

Paul Kump dell'Istituto di tecnologia dell'Illinois, ha sviluppato con il suo team un'altra strategia matematica per la lotta al crimine. In collaborazione con la Polizia di Chicago, ha analizzato la correlazione esistente fra i diversi crimini della città americana. Kump ha scoperto che il rischio che una persona compia un atto criminale cresce ogni volta che un soggetto sufficientemente vicino al suo "diagramma delle relazioni" compie anch'esso un reato. In altre parole, se una persona a me collegata delinque, la probabilità che anch'io agisca contro la legge aumenta significativamente per un periodo che i ricercatori hanno stimato in circa sei mesi.

Indicazioni come queste, ricavate dall'analisi dei network, sono preziose per le indagini. E non è un caso che anche i ricercatori di Messina siano oggi impegnati a "sviluppare delle strategie di intervento" più efficaci possibili, perché i pizzini si combattono anche con la matematica e l'informatica. (an.car.)