



Bei der Hamburger Polizei ist ein Analyse- und Darstellungstool im Einsatz, das zeigt, wo welche Verbrechen gehäuft auftreten. Den Berechnungen liegen verschiedene mathematische Ansätze zugrunde.

## Verbrechern auf der Spur

Seit dem Beginn der 1990er-Jahre sind vermehrt Geographische Informationssysteme (GIS) in die polizeilichen Arbeitsfelder vorgezogen, unter anderem auch in das Arbeitsgebiet der Lagebilderstellung oder Crime Mapping, wie es im englischen Sprachgebrauch heißt. Crime-Mapping wird unter anderem eingesetzt, um möglichst schnell Schwerpunkte in der Straftatenverteilung zu erkennen.

Seit dem Jahr 2001 besitzt auch die Hamburger Polizei mit dem System GISPOL eine Anwendung zur Straftatenanalyse. Diese Anwendung ist als Web-GIS-Applikation unter Nutzung der Esri-Produkte ArcIMS und ArcSDE eingerichtet und wird im Intranet der Hamburger Polizei betrieben. Ak-

tuell können bis zu 500 Nutzer gleichzeitig auf diese Anwendung zugreifen.

### Georeferenzierung

In der Regel wird ein Tatort als Punktobjekt in einer Karte dargestellt. In einem Geographischen Informationssystem ist dieser Punkt zu geokodieren, das heißt die Angabe der Adresse oder einer anderen Ortsangabe wird in in Koordinaten umgewandelt, die sich in der Karte darstellen lassen. Während bei einer Stecknadelkarte mehrere Nadeln an einem Punkt angebracht werden können und somit die Anzahl der Stecknadelköpfe die Anzahl der dort verübten Straftaten veranschaulicht, stellt sich die Darstellung mehrerer Objekte an einem einzigen Punkt in einem GIS schwieriger dar. Die verschiedenen Symbole liegen überein-

ander – die genaue Anzahl der Straftaten ist durch die reine Punktdarstellung nicht eindeutig erkennbar. Schwerpunkte in der räumlichen Verteilung von Straftaten sind auf diese Weise kaum zu ermitteln. Es zeigt sich also, dass die Punktdarstellung zur Identifikation räumlicher Muster nicht ausreicht. Es bedarf eines anderen Vorgehens, um die Schwerpunkte der Straftatenverteilung aufzudecken.

GISPOL stellt die Straftaten – nach Deliktarten getrennt – als Punktdaten in verschiedenen Layern dar. Die Straftaten werden jedoch nicht adressgenau dargestellt, sondern auf Straßenabschnitte bezogen, die vom Amt für Statistik für Hamburg und Schleswig-Holstein bereitgestellt werden. Diese Straßenabschnitte werden nach bestimmten Kriterien aus dem planaren Straßengraphen der Hansestadt

Hamburg erstellt. Sie stellen zum Beispiel Kreuzungspunkte und den Mittelpunkt eines Straßenabschnittes (Teil einer Straße zwischen zwei Kreuzungen) dar. Dieses Verfahren führt – insbesondere im Innenstadtbereich – häufig dazu, dass diverse Straftaten an einem Punkt übereinander liegen und erschwert somit das Erkennen von Schwerpunkten.

**Darstellung von Veränderungen**

Nicht nur die Analyse der räumlichen Verteilung von Straftaten, sondern auch die Kenntnis über die zeitliche Verteilung der Straftaten sowie die Identifikation von Schwerpunktverlagerungen im zeitlichen Verlauf ist von hoher Bedeutung für die Polizeiarbeit. Aus diesem Grund wurde ebenfalls versucht, die zeitliche Komponente mit in die Darstellung zu integrieren und auf diese Weise ein raum-zeitliches Abbild der Straftatenverteilung zu erhalten.

Um die nachfolgend beschriebenen Untersuchungen durchführen zu können, wurden von der Hamburger Polizei folgende Daten zur Verfügung gestellt

- Diebstahlereignisse des Jahres 2003
- Einbruchereignisse des Jahres 2003
- Straßenabschnittsdaten aus dem Jahr 2004

**Theoretisches Konzept**

Die theoretischen Grundlagen dieses Ansatzes entstammen der physikalischen Gravitationslehre. Das Potenzial wird nach der folgenden Formel berechnet:

$$P_i = \sum_{j=1}^n \frac{m_j}{d_{ij}^b}, d_{ii} > 0 \quad (1)$$

- P<sub>i</sub>: Potenzial des Ortes i,
- m<sub>j</sub>: Merkmalswert des Ortes j,
- d<sub>ij</sub>: Distanz zwischen den Orten i und j,
- b: Exponent der Distanz,
- n: Anzahl der betrachteten Orte

Einen Sonderfall stellt in dieser Gleichung die Berechnung des so genannten „Eigenpotenzials“ der Orte i dar (gemeint ist hier der Term m<sub>i</sub>/d<sub>ii</sub><sup>b</sup>). Da dieses Eigenpotenzial in die Berechnung des Potenzialwertes eingeht, ist Folgendes zu beachten: Die Orte i und j sind in diesem Fall identisch und somit ist der Distanzwert = 0. Blicke es bei diesem Wert, wäre die Gleichung nicht lösbar. Aus diesem Grund wird für d<sub>ii</sub> ein anderer Wert gewählt, der ≠ 1 ist. Sehr oft wird d<sub>ii</sub> = 1 gesetzt. Auf diese Weise entspricht das Eigenpotenzial

des Ortes dem Merkmalswert m<sub>i</sub> des Ortes i. Diese Behandlung des Eigenpotenzials führt jedoch zu einer Überbewertung des Eigenpotenzials im Verhältnis zu den übrigen Werten. Aus diesem Grund werden in der Literatur auch andere Ansätze zur Berechnung des Eigenpotenzials beschrieben.

Für die Hamburger Datensätze wurden zwei Varianten zur Berechnung der Potenziale eingesetzt. In der Variante 1 wurden die Potenziale für jeden einzelnen Straßenabschnitt ermittelt. Die flächenhafte Darstellung der Ergebnisse wurde anschließend auf Thiessen-Polygone bezogen, die aus der Verteilung der Straßenabschnittspunkte ermittelt wurden. Nun kam jedoch das im vorhergehenden Kapitel bereits erwähnte Problem der Eigenpotenziale zum Tragen. Um deren Einfluss möglichst gering zu halten, wurde ein bestehendes Verfahren zur Feststellung des Eigenpotenzials folgendermaßen modifiziert:

- Das zu jedem Straßenabschnitt gehörende Thiessen-Polygon wird gesucht.
- 2. Die Fläche des Polygons wird ermittelt und einer Kreisfläche gleichgesetzt.
- 3. Aus dieser Fläche wird der Radius des Kreises ermittelt.
- 4. Der Radius wird nun als Distanzfaktor d<sub>ii</sub> in den Term m<sub>i</sub>/d<sub>ii</sub><sup>b</sup> eingesetzt.

In der zweiten Variante wurde über die Untersuchungsfläche ein gleichmäßiges Raster gelegt. Anschließend wurde das Potenzial für jede dieser Rasterzellen entsprechend der oben vorgestellten Formel berechnet. Da die Wahrscheinlichkeit gering ist, dass der Mittelpunkt einer Rasterzelle mit einem Straßenabschnittspunkt zusammenfällt, entfällt das Problem des Eigenpotenzi-

als. Stattdessen trat ein neues Problem zutage. Die Entfernungen zwischen einer Rasterzelle und einem Straßenabschnittspunkt kann zufällig sehr klein werden, was dann bedeutet, dass im Rahmen der Potenzialermittlung der Wert dieser Rasterzelle (Anzahl der Straftaten an diesem Straßenabschnitt) überhöht wird.

Die Ergebnisse der beiden Verfahren sind in den Abbildungen 1 und 2 für die Verteilung der Einbruchdaten des Jahres 2003 dargestellt. Beide Darstellungen kommen zu annähernd identischen Verteilungen der Schwerpunkte. Insbesondere die Verteilung der höchsten Potenzialwerte ist in beiden Abbildungen annähernd identisch.

Die Rasterdarstellung weist tatsächlich einzelne Zellen mit extrem hohen Potenzialwerten auf, die in der ver- ▶

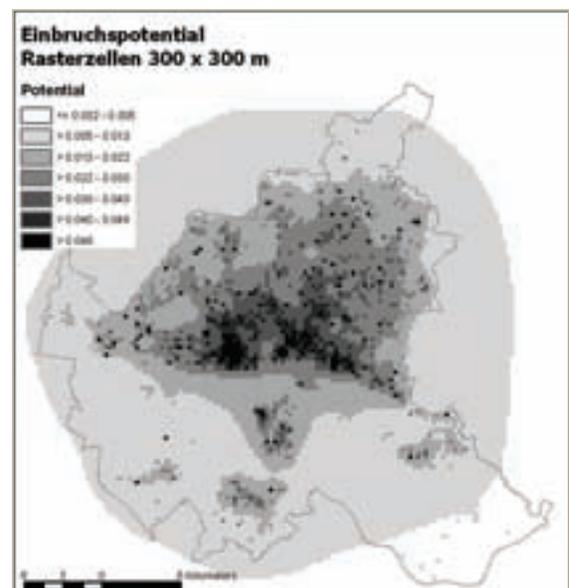
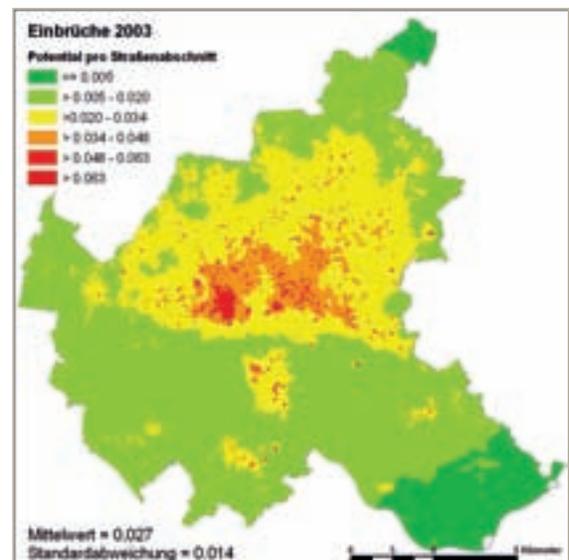


Abbildung 1: Darstellung des Einbruchspotenzials bezogen auf die Straßenabschnitte.

Abbildung 2: Darstellung des Einbruchspotenzials bezogen auf ein Raster.

gleichenden Darstellung der Straßenabschnitte nicht festgestellt wurden. Hier ist zu vermuten, dass diese Werte aus der zufälligen Nähe zwischen Rasterzelle und Straßenabschnittspunkt resultieren, was bei der Interpretation des Rasterbildes zu

### Raumzeitliche Verteilung

Ein weiteres Ziel dieser Arbeit bestand darin, die räumliche Verteilung der Straftaten vor dem Hintergrund zeitlicher Veränderungen aufzudecken und kenntlich zu machen. Zu diesem Zweck wurde die Diebstahlverteilung des Jahres 2003 im Wochenverlauf untersucht. Diese Untersuchung wurde sowohl mit Hilfe der Potenzialanalyse als auch mithilfe der Kernel-Density-Estimation vorgenommen. Um die Ergebnisse der Berechnungen anschließend darstellen zu können, mussten jedoch unterschiedliche Klassifizierungsverfahren herangezogen werden. So wurden für die Ergebnisse der Potenzialanalyse das Klassifizierungsverfahren der Standardabweichung und für die Ergebnisse der Kernel-Density Estimation die Klassifizierungsmethode der gleichen Intervalle herangezogen.

Ein Vergleich dieser beiden Methoden, der vor dem Hintergrund stattfindet, dass die Ergebnisse einem potenziellen Anwender als Animation vorgespielt werden könnten, lässt die Ergebnisdarstellung der Potenzialanalyse als wenig geeignet erscheinen, um dem Anwender räumliche Veränderungen zu präsentieren. Die hohen Schwankungen in den Klassengrenzen und -breiten, die aufgrund schwankender Mittelwerte und Standardabweichungen sowie durch Ausreißer in den Ergebnissen zu erklären sind, können die Anwender zu Fehlinterpretationen verleiten. Die Kernel-Density-Estimation zeigt sich hier besser geeignet, die zeitlichen Veränderungen interpretierbar darzustellen, denn durch das Klassifizierungsverfahren der „gleichen Intervalle“ verändern sich diese nicht von einer Abbildung zur nächsten.

Dennoch sind in beiden Bildserien zwei deutlich erkennbare Schwerpunkte der Diebstahlverteilung auszumachen. Diese wurden als das innerstädtische Einkaufszentrum zwischen Rathaus und Hauptbahnhof (der östliche Schwerpunkt) sowie das Vergnügungsviertel auf St. Pauli identifiziert (westlich). Die Potenzialanalyse zeigt zusätzlich zu diesen beiden

Schwerpunkten an fast allen Wochentagen einen noch weiter westlich gelegenen Schwerpunkt auf.

Erhebliche Unterschiede in der Aussage der beiden Darstellungen zeigen sich für den Sonntag: Während die Kernel-Density Estimation nur noch einen kleinen westlichen Schwerpunkt erkennt, zeigt die Potenzialanalyse einen großen Schwerpunkt, der sich über den gesamten Innenstadtbereich westlich und südwestlich der Alster bis zur Elbe hin erstreckt.

### Methodenbewertung

In die Bewertung der vorgestellten Methoden muss auch der jeweils erforderliche Zeitbedarf bis zur Ergebnisdarstellung einfließen. Da die Ergebnisse anschließend auch möglichst jedem Anwender bereit stehen sollen, sind als weitere Bewertungskriterien die Menge der durch den Benutzer vorzugebenden Berechnungsparameter sowie die Interpretierbarkeit der Ergebnisdarstellungen zu berücksichtigen.

Die Potenzialanalyse wurde zur Darstellung der räumlichen Verteilungen als hilfreich angesehen. Auch die Tatsache, dass keine Parameter einzugeben sind, sprechen für den Einsatz dieser Methode in der räumlichen Analyse von Straftatereignissen. Jedoch steht der hohe Zeitbedarf des in dieser Arbeit eingesetzten Berechnungsalgorithmus diesem Ansinnen entgegen. Sollte das Verfahren der Potenzialanalyse dennoch zur Darstellung von räumlichen Schwerpunkten der Straftatenverteilung eingesetzt werden, so muss der aktuelle Zeitbedarf zum Beispiel durch Optimierung des eingesetzten Berechnungsalgorithmus reduziert werden. Eine andere Überlegung könnte darin bestehen, dass die Möglichkeiten zur Formulierung von Berechnungsanfragen eingeschränkt oder auf die unmittelbare Zurücklieferung der Rechenergebnisse verzichtet wird. In diesem Fall könnte dann zum Beispiel ein Anwender per E-Mail einen Link auf das erstellte Rechenergebnis erhalten, sobald dieses erstellt worden ist.

Wie bereits erläutert wurde, eignet sich das Verfahren der Potenzialanalyse weniger, wenn es darum geht, in einer Bilderserie räumliche Veränderungen über einen definierten Zeitraum darzu-

stellen. Dies hängt mit der statistischen Verteilung der errechneten Potenziale zusammen, die in der Regel einer L-Verteilung folgen. In dieser Arbeit wurde zur Darstellung der Potenziale stets das Klassifizierungsverfahren der Standardabweichung genutzt. Dieses eignet sich jedoch nicht zur vergleichenden Darstellung von Bildsequenzen. Hier ist das Verfahren der Kernel-Density Estimation (welches die Klassifizierung mit Hilfe gleicher Intervalle ermöglicht) im Vorteil. Die Bildsequenzen sind leichter miteinander zu vergleichen, da die Klassengrenzen nur von dem Wertebereich zwischen dem kleinsten und dem größten Wert beeinflusst werden.

### Ausblick

Neben den hier vorgestellten Verfahren der Potenzialanalyse und Kernel-Density-Estimation wurden weitere Methoden untersucht, die in der neueren Zeit bereits in der räumlichen Straftatenanalysen zum Einsatz kamen. Hierbei handelt es sich um die Methoden der LISA-Statistiken. Diese Methoden, insbesondere der Gi\*-Index, weisen Regionen aus, in denen ähnliche Werte in unmittelbarer Nachbarschaft zueinander liegen. Diese Methoden sollten weiter untersucht werden.

#### AUTORIN:

Gabriele Dall  
T-Systems GEI GmbH  
Industry Business Line Public & Healthcare  
Service Delivery Center Homeland Security  
Lademannbogen 21-23  
22339 Hamburg  
E-Mail: Gabriele.Dall@T-Systems.com

