The Models on the Opioid Crisis

Ting Zheng *, Shangze Li, Guoqing Ji, and Genchao Zhang

College of Computer and information, Inner Mongolia Medical University, Huhhot, China
* Correspondence: ztyks@163.com

Abstract: The United States is experiencing a national crisis regarding the use of synthetic and non- synthetic opioids, either for the treatment and management of pain (legal, prescription use) or for recreational purposes (illegal, non-prescription use). In this paper, using the NFLIS data provided, build a mathematical model to describe the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time. Using the model, identify any possible locations where specific opioid use might have started in each of the five states.

Keywords: the Opioid Crisis; BP neural network; Spreading model

1. Introduction

The DEA/National Forensic Laboratory Information System (NFLIS), as part of the Drug Enforcement Administration's (DEA) Office of Diversion Control, publishes a data-heavy annual report addressing "drug identification results and associated information from drug cases analyzed by federal, state, and local forensic laboratories." The database within NFLIS includes data from crime laboratories that handle over 88% of the nation's estimated 1.2 million annual state and local drug cases. For this problem, we focus on the individual counties located in five (5) U.S. states: Ohio, Kentucky, West Virginia, Virginia, and Tennessee. In the U.S., a county is the next lower level of government below each state that has taxation authority.

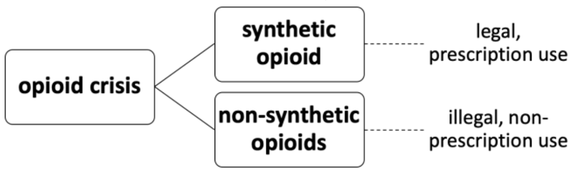

Figure 1. The opioid crisis

Supplied with this problem description are several data sets for your use. The first file (MCM_NFLIS_Data. xlsx) contains drug identification counts in years 2010-2017 for narcotic analgesics (synthetic opioids) and heroin in each of the counties from these five states as reported to the DEA by crime laboratories throughout each state. A drug identification occurs when evidence is submitted to crime laboratories by law enforcement agencies as part of a criminal investigation and the laboratory's forensic scientists test the evidence. Typically, when law enforcement organizations submit these samples, they provide location data (county) with their incident reports. When evidence is submitted to a crime laboratory and this location data is not provided, the crime laboratory uses the location of the city/county/state investigating law enforcement organization that submitted the case. For the purposes of this problem, you may assume that the county location data are correct as provided.

The additional seven (7) files are zipped folders containing extracts from the U.S. Census Bureau that represent a common set of socio-economic factors collected for the counties of these five states during each of the years 2010-2016 (ACS_xx_5YR_DP02.zip).

2. Models

2.1. Nomenclatures

i	an individual
Θ_i	geo-atom
Λ_i	geo-semantics
(S_i)	a point location
(T_i)	a given time
$\langle S, T, \Lambda_{ST} \rangle$	a qualified set of geo-semantics Λ_{ST} over space S and time T of interest
S	space
T	time
ID	unique identifier

2.2. Assumptions

BP neural network has strong non-linear processing ability. It can solve the problems of unclear background knowledge, unclear reasoning rules and complex information. It has unique advantages.Based on this, this paper establishes a hybrid model based on BP neural network and time-space series, and then takes settlement monitoring as an example to analyze and verify the practicability of the model.

The NFLIS data set provided was extracted from a much larger data set involving the five (5) states noted using "Narcotic Analgesics" and "Heroin" reports as the target. The variable Drug Reports indicates the number of identified drug cases corresponding to these target search topics in each county of each state. The variable Total Drug Reports County shows the total number of ALL identified drug cases in a county, of which 'Narcotic Analgesics' and 'Heroin' are a part, if they occurred. Because 'Narcotic Analgesics' and 'Heroin' are only two of many types of controlled substances, the sum of Drug Reports will not necessarily equal Total Drug Reports County for a specific county in a state for each reported year. Similarly, the variable Total Drug Reports State indicates the total of all drug reports for the state noted. However, since the database query was targeting 'Narcotic Analgesics' and 'Heroin' only, if a county had no such identified drug cases but had numerous drug cases of other types, the county will not appear in this dataset. Thus, the sum of all Total Drug Reports County for all counties in a state shown in this dataset for a particular year will not necessarily equal Total Drug Reports State.

2.3. The Foundation of Model

(1) Parameters

- i: an individual
- Θ_i: geo-atom
- Λ_i: geo-semantics
- (S_i): a point location
- (T_i): a given time
- $\langle S, T, \Lambda_{ST} \rangle$: a qualified set of geo-semantics Λ_{ST} over space S and time T of interest
- S: space
- T: time
- ID: unique identifier

(2) Geo-atoms $\Theta_i = \{S_i, T_i, \Lambda_i\}$

All geographic information can be decomposed into point sets or geo-atoms. An individual (i) geo-atom (Θ_i) consists of geo-semantics (Λ_i) measured, observed, or inferred at a point location (S_i) and a given time (T_i). Points are used here to refer to a location within which a piece of geographic information can be associated; hence, cells are a specific type of point. We consider this atomic form (i.e. geo-atoms) of geo-semantics at a location at a given time the most primitive unit of geographic information; all other types of geographic information consist of aggregations of the atomic form, often over infinite point sets. Our premise is that points are the most primitive spatiotemporal element with which information can be associated to space and time. Higher-dimensional properties of lines, areas, and volumes are aggregates of point sets by thresholding geo-semantics or assuming uniform geo-semantics in the aggregation. For example, property lots are aggregates of point sets that have the same ownership, contours are aggregates of point sets based on elevation criteria, and the State of Hawaii is an aggregation of a point set under the administrative authority of the State of Hawaii. Direct measurements of lines (e.g. distance or length) are not considered as geo-semantics at individual locations but are relationships (in the case of distance) between location points or secondary properties (in the case of length) by identifying all linearly aligned location points that satisfy a given geo-semantic threshold.

Geo-objects $\Omega = \{ID, S, T, \Lambda_{ST}\} = aggregate\ (\Theta | S, T)$

Geo-objects are what we identify as individuals in geography that cannot be further divided into individuals of the same kind. We consider that a geo-object is a uniquely identified (indicated by ID) location point set (S) and time set (T) in which geo-semantics meet certain requirements; $\langle S, T, \Lambda_{ST} \rangle$ indicates a qualified set of geo-semantics Λ_{ST} over space S and time T of interest. Notation (S) differs from notation (S_i) in that S denotes a point set whereas S_i marks an individual point. The same convention is applied to other notations throughout the paper. Under the consideration that geo-atoms are the most primitive units of geographic information, a geo-object is a function that aggregates geo-atoms (Θ) within space (S) and time (T) of interest. Locations of the point set may be represented by Cartesian coordinates, relative coordinates, or mathematical expressions (e.g., circles, arcs of ellipses, and Bézier curves) and may include disjoint subsets for geo-objects that consist of multiple parts (e.g., multipoints, multipart polylines, multipart polygons as recognized in the OGC Simple Feature Specification, www.opengeospatial.org).

Each geo-object must have a unique identifier (ID) to distinguish itself from the others. In some cases, the spatial location and extent of a geo-object are defined before geo-semantics are measured. Examples are census enumeration zones and lakes. Most current GIS data models take a space-centered approach that recognizes geographic objects by location and ascribes geometry to these objects. In doing so, new object identities are needed when changes occur to location or the geometry of an existing object. Alternatively, the formation of census enumeration zones can be considered as an outcome of a spatial aggregation based on geo-semantics which have been measured previously (e.g., previous census) or are easily distinguished without measurements (e.g., locations of water or not water). From this perspective, a geo-object identity is not determined by location or geometry, but by its intrinsic properties that make it a geo-object of its kind as judged by geo-semantic requirements.

In addition to properties assumed uniform over the object (spatially intensive properties), properties at the set level are likely to emerge and may include measures of the point set (e.g., length, area) and integrals of spatially intensive properties. These set measures and integrals are spatially extensive properties which are closely dependent on and cannot be separated from line or area objects. The contrast between spatially intensive and extensive variables is described in [1]**.** For example, population counts are a function of the area of the reporting zone. When a county is subdivided into two smaller units, the population density (a spatially intensive property) in its subdivisions may remain the same as the county population density, but population counts (a spatially extensive property) are likely to be different (unless one of the subdivisions has no population). Furthermore, many geo-objects have properties that are transitional in space. Their identities cannot be determined through aggregation of the geo-atoms that result from simply geo-semantic thresholding. Such a geo-object is characterized by its indeterminate boundaries and will be conceptualized here as a fuzzy point set [2].

When time is a consideration, the identities of geo-objects are critical to track through changes in location, geometry, and properties [3]. Since a geo-object is an aggregate of geo-atoms under a set of geo-semantic criteria, its identity is determined by the defined geo-semantic criteria and spatial and temporal constraints to the aggregation. The defined geo-semantic criteria specify the range or discrete values (e.g. domain) within which geo-semantics can be considered as the properties of a geo-object. Only when a geo-atom has geo-semantic values at $\langle S_i, T_i \rangle$ are within the domain, the geo-atom is part of the geo-object. When geo-semantic values meet the geo-semantic criteria at $\langle S_i, T_i \rangle$ but not at $\langle S_i, T_{i+1} \rangle$, the geo-object ceases to exist at S_i either by moving out of the location, dissipating entirely, or transforming into another geo-object with a different identity.

Once the identity of a geo-object is determined, its spatiotemporal path and behavior can be represented and tracked by lifelines [4]; and its spatiotemporal domain of accessibility can be represented by space-time prisms [5]. For geo-objects with geometry of higher dimensions (e.g. lines or polygons) or with multiple parts, space-time volumes are necessary to represent their spatiotemporal extent as elaborated in the SPAN ontology [6]. While observations of a lifeline or spatiotemporal geo-object are discrete, various interpolation methods (e.g. linear or curvilinear) may be applied to estimate intermediate locations and geometry between temporal observations. In addition to trajectories, additional parameters are necessary to record geo-objects which may change geometry over time. An example is the helix representation that uses a spline to track the location of a geo-object's centroid and prongs to record the extension of the geo-object in different directions at each point in time [7].

Aggregation of geo-atoms to form a geo-object is subject to spatial and temporal constraints by the nature of the geo-object, and such spatiotemporal constraints can be used to select appropriate interpolation methods as discussed above. A geo-object only exists in certain spatial and temporal extents bound by biological, physical, or administrative processes through which geographic entities are formed. At the highest level, no geo-objects on Earth can have a spatial extent greater than the surface of the Earth. Under constraints of physical processes, the largest hurricane recorded (Typhoon Tip) extended out to 1,100 km, and the smallest (Cyclone Tracy) was about 50km in radius. In addition, geo-objects have life expectancy; some may be ephemeral (e.g., rainstorms), but others can be long-lasting (e.g. mountains). Some geo-objects must be conterminous in space and time (e.g., a reservoir or a pollution plume), but others may have spatially or temporally disjoint parts (e.g., a wildfire or a country).

In summary, geo-objects are formed by aggregating geo-atoms under spatial, temporal, and geo-semantic constraints. Identities of geo-objects are recognized by spatial and temporal extents in meeting certain geo-semantic requirements. Changes to a geo-object over time can be tracked based on identities. Some geo-objects may have spatially or temporally disjoint parts. These geo-objects consist of discrete point sets $\langle S, T, \Lambda_{ST} \rangle$, but these discrete point sets are united by a common geo-object identity. On the one hand, geo-objects are 'discovered" by spatiotemporal aggregation of locations with qualified geo-semantics; in other words, geo-semantics are measured prior to the identification of geo-objects. On the other hand, geo-objects may be recognized by distinct geo-semantic discontinuity which enforces the perception of boundaries, and consequently the extents over which spatial and temporal aggregation takes place. In such cases, geo-semantics are characterized after the identification of geo-objects.

2.4. Solution and Result

Using the NFLIS data provided, build a mathematical model to describe the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time. Using the model, identify any possible locations where specific opioid use might have started in each of the five states.

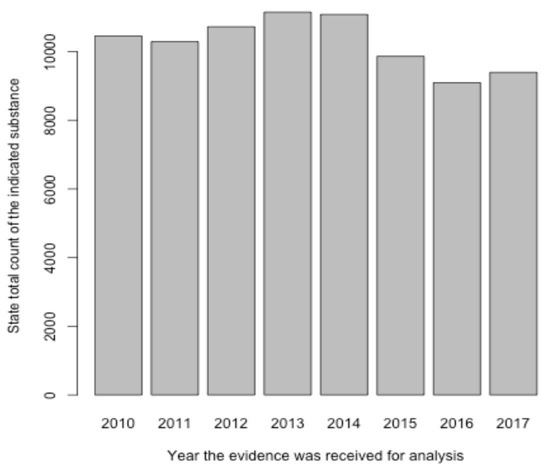

Figure 2. County total count of the indicated substance in Kentucky

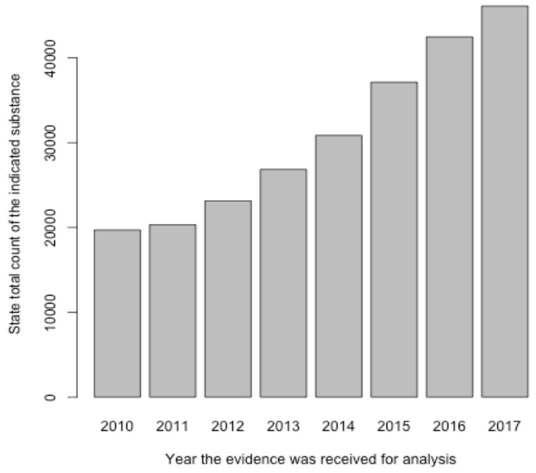

Figure 3. County total count of the indicated substance in Ohio

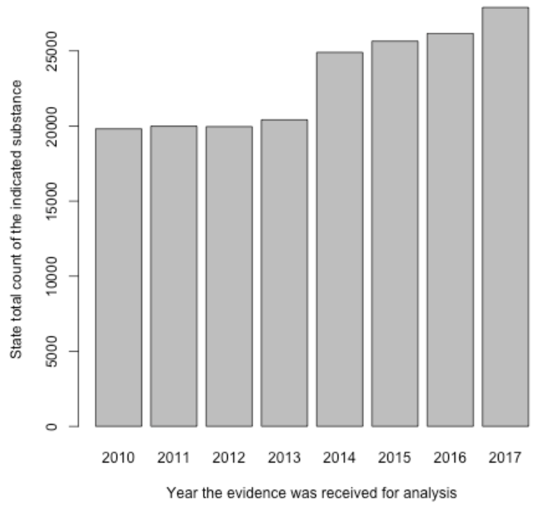

Figure 4. County total count of the indicated substance in Tennessee

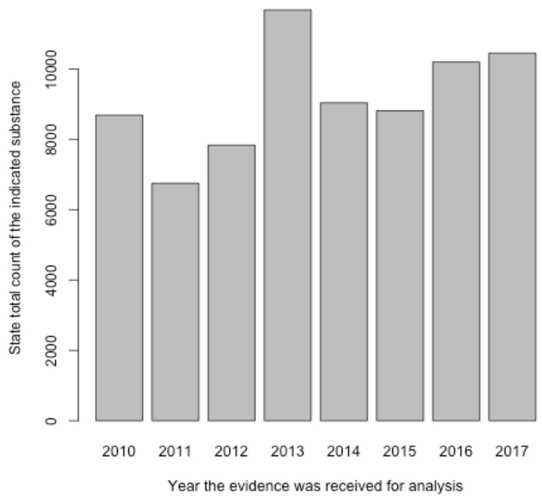

Figure 5. County total count of the indicated substance in Virginia

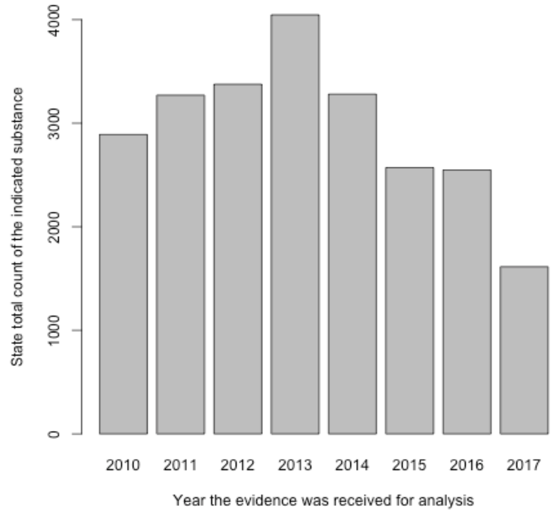

Figure 6. County total count of the indicated substance in West Virginia

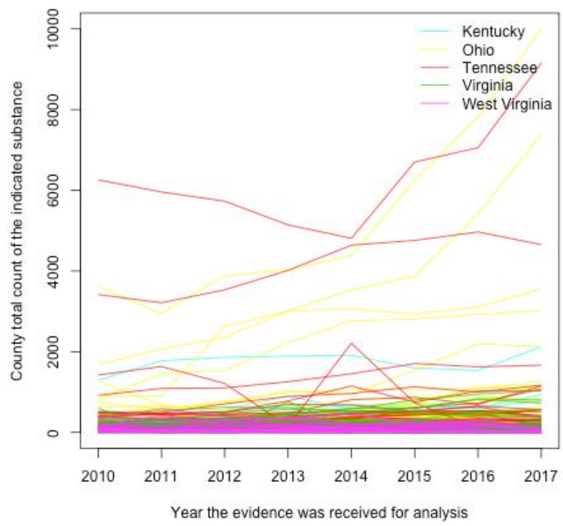

Figure 7. County total count of the indicated substance

We predict the location and time they will occur as follows.

Location (County, State)	Time (Year)
GUERNSEY, OH	2018
GREEN, KY	2018
EDMONSON, KY	2019
CLARK, OH	2018

3. Strength and Weakness

Strength: Most current GIS data models take a space-centered approach that recognizes geographic objects by location and ascribes geometry to these objects. In doing so, new object identities are needed when changes occur to location or the geometry of an existing object.

Weakness: A geo-object identity is not determined by location or geometry, but by its intrinsic properties that make it a geo-object of its kind as judged by geo-semantic requirements. “Portage” for example, was used as the county name for two spatially disjoint areas in Wisconsin at different times. When state-county names are used as county identifiers, the identity of Portage county is not tied to a particular geographic location.

Reference

[1] Agouris, P. and A. Stefanidis. Efficient Summarization of Spatio Temporal Events. *Communications of the ACM*. **2003**, 46(1): 65-66.

[2] Good child, M. F. Geographical data modeling. Computers & geosciences. **1992**, 18: 401-408.

[3] Hornsby, K. and M. Egenhofer . Identity-based change: a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*. **2000**, (14)3: 207-224.

[4] Cova, T. J. Extending geographical representation to include fields of spatial objects. *International Journal of Geographical Information Science*. **2002**, 16(6): 509-532.

[5] Miller, H.Modeling accessiblity using space-time prism concepts within a GIS. *International Journal of Geographical Information Systems*.**1991**, 5(3): 287-301.

[6] Good child, M. F. Geographical data modeling. Computers & geosciences. **1992**,18: 401-408.

[7] Grenon, P. and B. Smith. Spatial cognitiSNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial cognition and computation*. **2004**, (1): 69-104.

.