Nelson Foster | GEOG 6304 |Geographical Information Systems (GIS) I

GIS-based Research Proposal | Improving Missing Persons Data using GIS & Machine Learning

Abstract

Law enforcement personnel and community activists tend to lack comprehensive, consolidated, data-driven tools to address cases involving the "critically missing." According to District of Columbia (DC) Metropolitan Police Department's (MPD) missing persons website, a "critical missing person" is defined as any person under the age of 15 or over the age of 65, or anyone that, based on the specific circumstances (e.g., mentally incapacitated, patient who presents an imminent danger to him/herself or others, in a life threatening situation, real or suspected danger of foul play, etc.), is designated as such by the Patrol District's Watch Commander." In an effort to improve public communications regarding missing persons and hopefully expedite safe return, in 2017 the DC MPD instituted a practice of posting on Twitter whenever an individual is reported as "Critical Missing." As with any new effort to publicize information on crime or other social problems, it can create the perception of an upward trend or surge of such issues. A lack of national consolidated data for comparative context can add to the confusion.

This instance was no exception. As a result of the new practice, a group of activists, in an apparent miscalculation of the reported figures, alleged that fourteen black girls went missing in a twenty-four hour period. Shortly thereafter, several major news outlets ran stories about there being a disproportionate number of missing women in the DC metropolitan area, in particular black women. Due in part to the historical tendency for crimes and afflicting underrepresented sectors of the population to go either unreported and underreported, the story gained the attention of the Congressional Black Caucus and several other influential officials and groups. Under the hashtag and banner "#MISSINGDCGIRLS, public discontent culminated in "dozens of community activists and parents — including a mother who said she had struggled to get help from the police when her 12-year-old child was missing for a week — gather[ing] outside the African American Civil War Memorial for a Protect Black Kids candlelight vigil" (Stolberg, 2017). After the proverbial dust settled, MPD later confirmed that "at no point…ha[d] 14 girls disappeared from the city in a single day" (Dwyer, 2017). However, "[a]ctivists argue the inaccuracy of the post itself should not detract from the wider issues it has highlighted: the dangers confronting runaway youth, and the racial dimensions of how law enforcement treats missing kids (Dwyer, 2017). This paper will review relevant research and current trends in the utilization of spatial modeling and GIS technologies to identify, track, and catalog data on missing persons, and propose a method to create a consolidated geodatabase to capture attribute and spatial data on missing persons, the features of which to be used in a predictive model to aid law enforcement in the timely and successful locating of the critically missing.

Literature Review

"Investigating missing person cases: how can we learn where they go or how far they travel?"

In this article, several lecturers and scholars in the fields of Criminal Justice and GIS at the University of Portsmouth delve into the spatial element of missing person's data in an effort to identify how it can improve policing policy in the United Kingdom. According to their research, aside from basic demographic information on the missing person (age, amount of time missing, etc), "there is very little information on where or how far missing individuals travel…[identifying such spatial behavior patterns (i.e. behavior in space) is highly important, as such information may be used by police forces to identify and refine search areas" (Shalev, Shaefer, & Morgan, 2009). This scarcity in spatial data is further complicated by the fact that at the time of the article, there was no consolidated missing person's database in the UK. Most data was captured by local police forces and data sharing between police forces and with advocacy groups was limited. Also, there was no consistent protocol on the type of information gathered from missing persons once they are recovered, or from individuals who were harboring them, which could have been critical in gathering the spatial behavior data the authors speak to. The research advocates for a "problem solving" approach that utilizes the same "crime mapping" approach for missing persons data. The authors conducted a pilot study using data collected on 577solved missing person cases during 2003. They focused on such spatial elements as location the person went missing, a location they were known to known to have visited while missing (although this was hard to corroborate and was not used in the analysis) and location where they were found. This data was geocoded via the UK's "Ordnance Survey Code-Point" data using Digimap, and was enhanced by demographic attribute data, after which spatial queries were conducted to determine distance traveled. The study yielded that the majority of people went missing from their home, however determining average distance proved difficult, as most individuals returned to the place where they went missing from. Though the study is highly constrained by missing data, the methods are very promising.

"Investigating missing person cases: how can we learn where they go or how far they travel?"

This article provides a discussion of highlights the importance of geospatial data and methods for locating and recovering missing persons in a wide geographical jurisdiction: New South Wales, Australia. To address this problem, "police forces are using powerful geographic information systems to facilitate precise, coordinated, collaborative land search and rescue operation (Berry, 2017). The article highlights the importance of "Last Known Position (LKP)" in establishing a search radius. After this is obtained or estimated, software from the security firm Pitney Bowes is used to collate data from various sources (witnesses, mobile phone signal areal, "Probability of Detection" calculations) to create maps for search teams to utilize, complete with buffers, rings, polygons, etc. to help the effort. Because of this software, the Constable of New South Wales asserts that search teams "can accurately coordinate a land search, send out the correct people and maintain positional accuracy" (Berry, 2017)/

"GIS for Missing Persons"

This article provides a case study regarding how individuals leveraged email, social media, "citizen forensics" and ArcMap to effectively "crowdsource" a search effort for a missing individual in the San Francisco Bay Area with "a mild disability which delayed his mental development "(Bishop, 2012). Though many individuals were willing to help, few knew where to begin. The family of the missing individual had a friend with GIS skills who was contacted to assist. According to the individual's banking records, they were able to find when their ATM card was used, narrowing the search, however the search radius proved to be in a dangerous area. Despite this, "ArcMap 10 desktop was utilized to create custom search maps with several layers of relevant data" (Bishop, 2012) In addition, a mobile map using the "Crowdmap" platform was created to help identify flyer locations, sightings, and other day. Unfortunately, the individual was not recovered alive. However, the lessons learned from the ordeal provided the author with a point of departure to develop a nonprofit organization to map missing person's data with the support of various third party entities.

Methodology

A common theme of a "lack of consistent data" can be found throughout the literature. The U.S. is very similar to the U.K. to extent, in that even some of the federally funded databases such as the National Missing and Unidentified Persons System are volunteer in nature. Funding and concerted effort is critical.

To address this issue, the following course of action is proposed:

Attribute Data Acquisition: Through direct download, Application Program Interface (API), or Freedom Information Act (FOIA) request, missing person's data will be gathered from nine major U.S. Municipalities and two federal databases, to include any available spatial data.

Attribute Data Preprocessing: Attribute data will be normalized and loaded into a geodatabase using ArcCatalog, with any XY coordinate data converted to spatial elements using the correct respective State Plane Coordinate System.

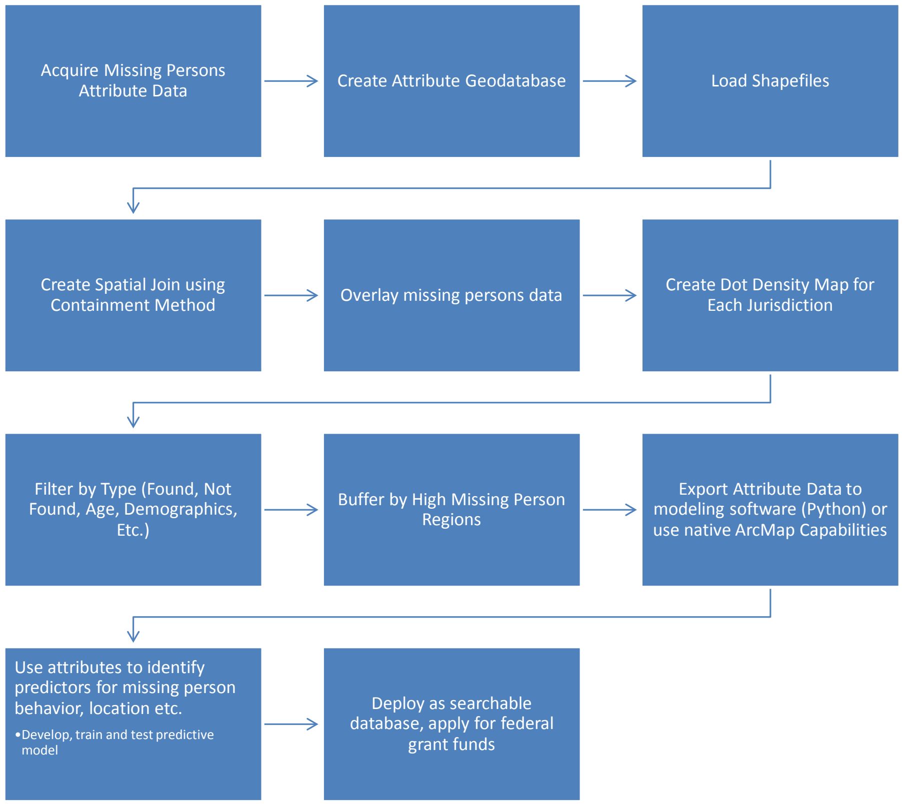
Spatial Data Acquisition: For each major metropolitan area, the local OpenGIS data repository will be leveraged to obtain necessary boundary, road centerline , hydrology, and other pertinent shape files in order to build necessary base maps.

Spatial Analysis: For each metropolitan area. Dot density maps will be created using the missing person data. Spatial joins will be used to identify various locations via adjacency, containment and other methods, and this data will be output as .dbf files for modeling.

Predictive Modeling: For the predictive model, I propose using the attribute table data generated from various spatial joins from the geodatabase as "features" in order to create testing and training datasets to inform a Spatial Regression Model utilizing the (Ordinary Least Squares) method in ArcGIS and or Python (GISGeography, 2017). The Model can help understand the variables and features that explain why vulnerable groups of people go missing, and help potentially predict their movements so that they can be found more quickly. Also, some of the predictive methods utilized by search and rescue operations in rural or remote areas could also be adapted this effort (Filipkowska, Koester, Chrustek, and Zarod, 2012).

Apply for Funding: Use the results of the aforementioned efforts in a Department of Justice Grant Proposal in order to hopefully get a formal national missing persons geospatial intelligence database created and deployed to local law enforcement and nonprofit organizations.

High-Level Methodology Flowchart



Geospatial Layer "Wish List"

1. District of Columbia Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://missing.dc.gov/
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
2. New York Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: http://www.criminaljustice.ny.gov/missing/
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)

3. Chicago Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://home.chicagopolice.org/category/missing-persons/
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
4. Los Angeles Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://oag.ca.gov/missing/search
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
5. Cleveland Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: http://missingpersonscuyahoga.org/en-US/AllMissingPersons.aspx
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
6. Miami Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://www.miamidade.gov/police/missing-persons.asp
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
7. Atlanta Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://gbi.georgia.gov/cases/missing-persons
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
8. Atlanta Missing Persons Data
    a. Type: Attribute
    b. Cost: Low/Medium
    c. Source: https://gbi.georgia.gov/cases/missing-persons
    d. Date/Year: most current data available
    e. Other relevant info (any xy or geocode-able data)
9. Basemap Data (Boundary, Road Centerline, Hydrology)
    a. Type: Shapefile
    b. Cost: none to low
    c. Source: OpenGIS Websites for Washington, DC, New York, Chicago, Los Angeles, Cleveland, Miami, and Atlanta, and/or https://catalog.data.gov

References

Bishop (2012). GIS for Missing Persons. *University of Redlands. Retrieved from https://www.redlands.edu/globalassets/depts/school-of-business/gisab/student-awards/2012_winning_submission_1_mapping_for_missing_persons1.pdf*

Barry, A. (2017, October 27). Using GIS to Help Locate Missing Persons. *GIS Lounge. Retrieved from https://www.gislounge.com/using-gis-help-locate-missing-persons/*

Dwyer, C. (2017, March 27). DC's Missing Teens: A False Number that Spurred a Real Conversation on Race. *The Two-way: Breaking News from National Public Radio (NPR). Retrieved from https://www.npr.org/sections/thetwo-way/2017/03/27/521655564/d-c-s-missing-teens-a-false-number-that-spurred-a-real-conversation-on-race*

Filipkowska, E. (Fall 2012). Modeling the Movement of Lost People. *ArcNews* Retrieved from http://www.esri.com/news/arcnews/fall12articles/modeling-the-movement-of-lost-people.html

Filipkowska, E., Koester, R.J., Chrustek, R., and Zarod, M. (Fall 2012). Lost and Found in the Polish Carpathian Mountains – GIS-Based analysis is Aiding the Search, Rescue, and Safe Return of Lost Hikers. A*rcNews.* Retrieved from: http://www.esri.com/news/arcnews/fall12articles/lost-and-found-in-the-polish-carpathian-mountains.html

How to Build Spatial Regression Models in ArcGIS. (2017, October 4). Retrieved from https://gisgeography.com/spatial-regression-models-arcgis/

Shalev, K., Shaefer, M., and Morgan, A. (2009). Investigating Missing Person Cases: How Can We Learn Where They Go or How Far They Travel? *International Journal of Police Science & Management, Volume 11, Number 2. Retrieved from http://journals.sagepub.com/doi/pdf/10.1350/ijps.2009.11.2.116*

Stolberg, S. G. (2017, March 31). Missing Girls in Washington DC Widen City's Racial Divide. *The New York Times. Retrieved from https://www.nytimes.com/2017/03/31/us/washington-dc-missing-girls-children-trafficking.html*