Safety-aware routing for motorized tourists based on Open Data and VGI

Andreas Keler, Jean D. Mazimpaka

Applied Geoinformatics, Institute of Geography,
University of Augsburg
Introduction (1)

• Problem statement (1)
  – As part of LBS, routing solutions have attracted a lot of interest
  – Most of these solutions aim at finding the shortest route
  – The need to consider personal preferences has led to personalised routing solutions (e.g. finding a best route through Points of interest)

Shortest route between two places in Boston [3]  
Beautiful route between 2 places in Boston [3]
Introduction (2)

• Problem statement (2)

  – Safety is important in routing, especially for categories of people such as tourists travelling during the night in unfamiliar environment

  – However, few studies have considered safety for routing and done so implicitly

  – Furthermore, the potential of the increasingly available Open Data and VGI for safety consideration in routing has not been explored

• Objective

  – Integrating open data and VGI for a routing approach personalised for the class of people for which safety is the major routing constraint
1. Open governmental data:
   - Police stations (a)
   - Street lights (b)
   - Crime occurrence data
2. VGI:
   - Highway data (c), extracted from OSM,
     Road segments with key highway equals to motorway or primary,
     Converted into points
Investigation area and used datasets (3)

Safety weights ($w_A$, $w_B$, $w_C$) assigned to point objects of the 3 types of datasets

Low

Street light

Street light

Highway

Street light

Police station
Approach (1)

1. Defining a safety index and creating a cost surface

- Merging the 3 types of point data: Police station (A), Street light (B), Highway (C)

- For any point $P_i$, the safety index $I_{safe}$:

\[
I_{safe}(P_i) = \begin{cases} 
  w_A, & \text{if } P_i \in A \\
  w_B, & \text{if } P_i \in B \\
  w_C, & \text{if } P_i \in C \\
  \text{Interpolate}(w_A, w_B, w_C), & \text{otherwise}
\end{cases}
\]

- Inverse Distance Weighting (IDW) used as the point interpolation method
Approach (2)

- Introduce the danger index $Idang$ (b) as an inverse of $Isafe$ (a) $Idang = \text{inv}(Isafe)$
Approach (3)

2. Generating a weighted road network
   - The cost surface made of raster cells with danger index values
   - Averaging pixel values of raster cells that intersect a given road polyline

3. Deriving routing obstacles from crime data
   - A density-based clustering method used to build clusters of crime occurrence locations
   - Clusters converted to polygons through convex hull generation
Approach (4)

4. Combination of weighted road network and crime hotspots

- Polygons representing crime hotspots to be used as obstacles for routing

- Safest route computed as a least cost path with:
  
  - Cost based on danger index connected to road network
  
  - Crime hotspot polygons as obstacles to be avoided
Results (1)

- Approach tested on a case study of routing during the night for tourists driving without local knowledge

- Route proposed by the approach (c) compared to the one proposed by the commonly used approach (the shortest path) (a)

- As expected the two routes are very different
Results (2)

- Effect of considering further constraints seen by comparing the route avoiding crime hotspots (c) and the route that does not (b)
Conclusion

- A safety-aware routing developed based on the combination of open data and VGI
- The approach can provide insights into the potential of the ever increasing amount of freely available data for navigation purpose
- The approach can be extended by considering more safety factors
- Other directions of extending this work include:
  - Establishing a taxonomy of dangerous features and providing route alternatives based on it
  - Exploring the possibility of implementing a safety-aware routing integrating near real time data (e.g. from social media)
  - Exploring the potential of using historical crime hotspots for improving efficiency (e.g. on near real time data)
After LBS 2015 in Augsburg, a new commission of the International Cartographic Association (ICA) established:

**ICA Commission on Location Based Services**

http://lbs.icaci.org/

and will be published this year in Journal of Location Based Services

This work was presented last year at the
Thank you for your attention!
References


Notes (1)

Assumptions:

• A higher crime frequency in night time [1] 
  [a relationship exists between violent crime and alcohol; premises that serve alcohol 
on-site (bars, nightclubs) bear a considerable responsibility for this alcohol-fueled 
violence. Generally, these premises operate in the night]

• Police stations are assigned a higher safety weight because the closer to 
a police station a location is the lesser a crime/danger is likely to occur 
there

• Next higher weight assigned to Highway because we assume that the 
high speed used reduces the likelihood of a danger commonly associated 
with a stoppage (compared to lower level roads)

• Street lights receive a safety weight but lower than that of the first two 
classes because we assume that locations with street lights have lower 
risk of night crimes than locations without lights. [improved street 
lighting led to reductions in crime [4] ]
Common point interpolation methods:

- Kriging
- Linear point interpolation
- Spline interpolation
- Inverse distance weighting (IDW): each measured value has a weight that is inversely proportional to the distance to the estimated point values
IDW [2]

- A general form of finding an interpolated value $u$ at a given point $x$ based on samples $u_i = u(x_i)$ for $i=1,2, ..., N$ using IDW is an interpolating function:

$$u(x) = \begin{cases} 
\sum_{i=1}^{N} \frac{w_i(x)u_i}{N}, & \text{if } d(x, x_i) \neq 0 \text{ for all } i \\
\sum_{i=1}^{N} w_i(x), & \text{if } d(x, x_i) = 0 \text{ for some } i \\
u_i, & \text{if } d(x, x_i) = 0 \text{ for some } i 
\end{cases}$$

where

$$w_i(x) = \frac{1}{d(x, x_i)^p}$$

- In our case; The interpolation used a variable search radius, which is dependent on the number of input points.
- We used 100 input points. The resulting cell size of the raster is 100m.
OPTICS

• Density-based clustering with:
  
  – Distance threshold = 500m
  
  – Minimum number of objects = 3
Notes (5)

Data sources


• The highway data: [http://wiki.openstreetmap.org/wikiKey:highway](http://wiki.openstreetmap.org/wikiKey:highway)

• The Police station locations data were created on 24/7/2012

• The crime data come from the Los Angeles County Sheriff’s Department‘s jurisdiction ([http://www.lasd.org](http://www.lasd.org))
  – The crime data are extracted from the reports of the last 30 days (in our case reports covering the period 19/1 – 17/2/2015)
  – We considered only Part I crimes, which include among others robbery, aggravated assault, burglary, larceny theft, grand theft
Case study:

- Routing from the Long Beach Airport (in the south of the study area) to the Hollywood Boulevard (in the north)

- Considering more factors (e.g. Venues that serve alcohol on-site, …) leading to:
  - more heavy computation
  - difference of the route section where obstacles are