Spatial Crime Patterns vs Safety Perception: Mixed Experiments

Alina Ristea

(1) Postdoctoral Research Associate, (2) Research Fellow

(1) Boston Area Research Initiative (BARI)
School of Public Policy and Urban Affairs
Northeastern University, Boston

(2) National Center for Spectator Sports Safety and Security
The University of Southern Mississippi
About me

-> Bachelor in Geography and Cartography;

-> Master in Geographic Information Systems (GIS), Department of Geography, University of Bucharest, Romania

-> PhD in Applied Geoinformatics, University of Salzburg, Austria: *Integration and evaluation of social media in crime prediction models*

-> Postdoc at Boston Area Research Initiative, CSSH, Northeastern University

Interests: spatial crime analysis, social media mining, predictive analytics, safety perception, neighborhood analysis, GIScience, urban informatics
Outline

PART 1: Social media in crime prediction models

● Background
● Scope (research gaps and objectives);
● Results
  ○ Research Objective 1;
  ○ Research Objective 2.
● Discussion;
● Scientific contribution;
● Future directions of research;

PART 2: Applying Geospatial Technology to Explore Urban Blight and Perceived Safety

● Background
● Scope
● Data and Geospatial Technologies;
● Results;
● Relevance;
● Future work.
Thesis Context

**Predictive policing** -> goal of preventing crime, solving past crimes, and identifying potential offenders and victims. (Perry et al. 2013)

**Social media mining** -> process to extract patterns, form conclusions about users, and act upon the information, often for the purpose of advertising to users or conducting research. (Zafarani et al. 2014)
Environmental criminology

important theoretical foundation for exploring spatial crime distribution (Bruinsma and Johnson, 2018).

Routine Activity Theory 1979

Crime Pattern Theory 1981

Key concepts: crime attractors, generators and detractors
Which features can we extract from social media?

Kurland, Tilley and Johnson (2014)

Botta, Moat and Preis (2015)

Gerber (2014)

Struse and Montolio (2014)

Research gap

RO 1. Space and time relationships

RO 2. Space and time crime prediction

Sporting events

Crime occurrences

Social media data

Improved prediction?
Objectives

Research Objective (RO) 1: Uncover relationships between crime patterns and social media posts

RQ1: Does social media activity (i.e., tweets) correlate in space and time with crime occurrences?

RQ2: Do different crime types show distinct relationships with tweet–related features?

RQ3: Does the distribution of social media posts follow the changes in urban crime patterns when a sporting event occurs?
Objectives

**Research Objective (RO) 2**: Improve methods for integrating social media data into crime prediction models

RQ4: Do geo-located tweets improve crime prediction models and enrich the information coming from historical crime data and additional explanatory variables?

RQ5: Can tweets be a factor for determining at-risk populations?

RQ6: Does the use of social media as a dynamic feature have a higher relevance in prediction models related to non-routine activities, rather than ordinary ones?
Case studies

- Football and crime
- Crime spots patterns
- Hockey and crime
- Population at crime risk
- Basketball, hockey, and crime
- Evaluation methods

United Center cleared after reports of fighting during basketball tournament
Data (pre) processing flow

Setting the scene
- study area
- crime types

Temporal filtering
- home/away game days vs control days
- 8-hours time frame; 24-hours time frame

Independent variables
- environmental, demographic, socio-economic
- Twitter-related features

Methods
- Explanatory models
- Prediction models
Tweet-related features

Sentiment analysis
- polarity;
- emotions (feelings).

Topic modeling extraction
- sporting events;
- violent topics.

Population at crime risk
- ambient population;
- complement residential population.

Crime-related tweets
- lexicon based;
- mixed approaches.

Plutchick’s wheel of emotions

- optimism
- losiness
- joy
- serenity
- interest
- anticpation
- ecstasy
- trust
- vigilance
- acceptance
- tenor
- fear
- admiration
- love
- foreboding
- loathing
- disapproval
- appreception
- contempt
- catastrophe
- distraction
- disapproval
- disapproval
RO 1: Uncover relationships between crime patterns and social media posts

Methods: spatial autocorrelation Moran’s I, bivariate autocorrelation, Pearson correlation, density mapping, comparison between game days and control days, sentiment analysis, topic modeling, crime-related text extraction.
Density maps Aston Villa stadium (a) amalgamated crimes, (b) geotagged tweets, (c) violent tweets, and (d) football-related tweets
Density (a) criminal damage, (b) theft and handling, and (c) violence against the person.
Crime occurrences

Twitter posts

Hours Before and After Kick-off (KO)

Sunday crime intensity (%)

Sunday tweets intensity (%)

RO 1 -> RQ3

P2
Bivariate LISA clusters between crime density and (a) tweets density, (b) violent tweets density, and (c) football topic tweets density.

Legend - Cluster Map:
- Not significant
- High-Low
- Low-High
- Low-Low
- High-High

- Stadium
- Grid size 500m
- Railways
- Roads

Game days vs. Control days: RO1 -> RQ1

(1a) (1b) (1c) (2a) (2b) (2c)
Crime density 3km area around Rogers Arena, Canucks team (hockey), Vancouver, Canada
Crime density 3km area around Rogers Arena, Canucks team (hockey), Vancouver, Canada
RO 2: Improve methods for integrating social media data into crime prediction models

**Methods:** Geographically Weighted Regression (GWR), Negative Binomial Regression (NBLR), Logistic Regression, Random Forest, (Local) Kernel Density Estimation (LKDE and KDE), density weighted areal interpolation.
Crime-tweets influence on theft-from-vehicle spatial distribution

Legend
- Stadium-Rogers Arena
- Coefficient crime-tweets
  - ≤ 0
  - 0.001 - 0.050
  - 0.051 - 0.100
  - 0.101 - 0.150
  - 0.151 - 0.200
  - 0.201 - 0.250
  - > 0.250
- Dissemination areas (DA)

Home games and comparison days
Away games and comparison days
GWR models
Mischief Adj. R² | Game days | Comparison days
--- | --- | ---
Tweets | 0.83 | 0.78
Anticip. | 0.85 | 0.79
Surprise | 0.82 | 0.78
Trust | 0.85 | 0.78
Positive | 0.84 | 0.79
Predicting crime types

RO 2 -> RQ6

Assault - AUC improvement

Robbery - improvement AUC

Home games Bulls  |  Away games Bulls  |  Home games Blackhawks  |  Away games Blackhawks  |  Control days
Game days vs control days prediction

RO 2 -> RQ6

Home games Chicago Bulls – AUC improvement

Control days – AUC improvement
Controversial result

**RO 2 -> RQ6**

---

**Home games Chicago Blackhawks – AUC improvement**

- **Assault**
- **Battery**
- **Criminal damage**
- **Motor Vehicle Theft**
- **Other offense**
- **Robbery**
- **Theft**

Legend:
- **Additional**
- **Tweets and historical**
- **Tweets, historical and additional**
- **Violent and historical**
- **Violent, historical and additional**

---

**Home games Chicago Bulls – AUC improvement**

- **Assault**
- **Battery**
- **Criminal damage**
- **Motor Vehicle Theft**
- **Other offense**
- **Robbery**
- **Theft**

Legend:
- **Additional**
- **Tweets and historical**
- **Tweets, historical and additional**
- **Violent and historical**
- **Violent, historical and additional**

---
Density weighted areal interpolation technique

**RO 2 -> RQ5**

**Residential population** = consists on residents who permanently stay in an area for a considerable amount of time and are part of the official population count;

**Ambient population** = refers to the actual number of persons who are present within a particular area at any given time.
RO 2 -> RQ5

Mean values of Hit Rate

<table>
<thead>
<tr>
<th></th>
<th>cell size</th>
<th>length of prediction period</th>
<th>method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cell A*</td>
<td>1 week</td>
<td>Baseline</td>
</tr>
<tr>
<td></td>
<td>24.9</td>
<td>18.7</td>
<td>19.7</td>
</tr>
<tr>
<td></td>
<td>cell B</td>
<td>2 months</td>
<td>GWR*</td>
</tr>
<tr>
<td></td>
<td>20.0</td>
<td>23.0</td>
<td>25.2</td>
</tr>
<tr>
<td></td>
<td>3 months*</td>
<td>3 months*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Discussion

<table>
<thead>
<tr>
<th>Explaining relationships</th>
<th>Data characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event routine activity</td>
<td>Data quality: social media bias and geo-location</td>
</tr>
<tr>
<td>Need of control variables</td>
<td>Geo-privacy for crime data</td>
</tr>
<tr>
<td>Fan behavior</td>
<td>MAUP and temporal unit selection</td>
</tr>
<tr>
<td>Significant crime-crime tweets relationship</td>
<td>Data sparsity: negative-positive ratio</td>
</tr>
<tr>
<td>Population at crime risk</td>
<td>Transferability</td>
</tr>
<tr>
<td>Prediction day vs training data</td>
<td>Differences per crime types/culture/country</td>
</tr>
</tbody>
</table>
Scientific contribution

• emerging field of *predictive analytics*;

• *geography of crime* for sporting events;

• collaboration based - highly *interdisciplinary outcomes*;

• evaluating *significance of social media* features in prediction models;

• spatial *hot spots and cold spots* analysis;

• *text* analysis in the *space-time* view.
Future directions of research

- hot spots vs cold spots;
- ambient population;
- subjective safety perception vs objective crime.

- real time crime prediction;
- dynamic spatiotemporal features in prediction;
- testing novel spatiotemporal performance evaluation.

Applications: crime prevention strategies and law enforcement, policy makers, law enforcement, urban design for events, crime safety regulations, sports analytics.
Outline

PART 1: Social media in crime prediction models

- Background
- Scope (research gaps and objectives);
- Results
  - Research Objective 1;
  - Research Objective 2.
- Discussion;
- Scientific contribution;
- Future directions of research;

PART 2: Applying Geospatial Technology to Explore Urban Blight and Perceived Safety

- Background
- Scope
- Data and Geospatial Technologies;
- Results;
- Relevance;
- Future work.

Collaborators: Michael Leitner, Judith Stratman, Bernd Resch, Kalliopi Kyriakou
Broken Windows Theory

Kelling and Wilson (1982)

It states that “visible signs of crime, anti-social behavior, and civil disorder create an urban environment that encourages further crime and disorder, including serious crimes” ➔ great debate in criminology and not only!!!

Disorder goes untreated ➔ Citizens become fearful and withdraw from the community ➔ Informal social control decreases and/or is perceived to be low by criminals ➔ Disorder and crime increase as criminals increase their activity in the area

Broken windows effect (Hinkle and Weisburd 2008)
Motivation and goals

- Identifying physical urban blight indicators and find correlations with crime data;
- Applying new methods to observe urban neighborhood characteristics and to include qualitative data into a GIS;
- Extracting safety information from the data acquired using mixed methods and to implement it in a GIS.

As a long-term outcome, we would like to contribute to improving citizen’s cooperation with official stakeholders and help to design crime prevention strategies
Study area

East Baton Rouge Parish

Crime density in the city of Baton Rouge

- very low crime
- low crime
- moderate crime
- high crime
- very high crime

Swamps
Study area
Methodology. Pre-analysis for field work

- **Spatial unit selection**: neighborhood for field selection, Census blocks for interpretation;
- **Defining categories**: very high, high, moderate, low, very low crime rates;
- **Selection criteria**: no highway; no lakes; connectivity; similar length of street network
- Determining the shortest path for driving in the neighborhood;
Methodology. Data acquisition

Primary data collection:
• Survey: background questionnaire and on-screen mapping;
• Spatial video acquisition system (SVAS);
• Geonarratives;
• Physiological measurements using wristbands;

Secondary data collection:
• Crime data;
• Additional: socio-demographic and environmental data.
(1) Spatial video acquisition system (SVAS)

- additional technique to GIS to improve the documentation and analysis;
- unlike Google Street View, SVAS data collection is in the control of the researcher;
- spatial video can be collected using a variety of modes (car, motorbike, bicycle, boat and by foot);
(2) Geonarratives

- gives contextual details and enriches typical hotpot approaches with more on-the-ground context;
- audio recording of this narrative is linked to the video via timestamp;
- multiple perspectives for the same geographic area;
- mental map from behavioral geography.

Example of geo-narrative output
(3) Physiological measurements - wristbands

- tool for capturing people’s subconscious reactions to environmental stimuli;
- add contextualizing information to observed phenomena;
- can complement videos and narratives;

Source: www.empatica.com
(4) Crime, socio-economic, demographic data

- Baton Rouge Police Department (BRPD) including coordinates and time stamps of crime occurrences;

- Census data: residential population, ethnicity, education, household types, foreign born, unemployment, poverty rate;

- Environmental data: street network, buildings footprint, public buildings, neighborhoods, etc.
Methodology. Data processing

Processing tools:
- Videoplayer with integrated GPS track and
- WordMapper (developed by Prof. Andrew Curtis and his team)

Analysis tools:
- R programming for statistical analysis
- GIS software for mapping
Results

Urban blight > Safety perception
Physical Blight – data extraction

Building

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>Abandoned or boarded up properties</td>
</tr>
<tr>
<td>B2</td>
<td>Broken window</td>
</tr>
<tr>
<td>B3</td>
<td>Blocked window</td>
</tr>
<tr>
<td>B4</td>
<td>No window</td>
</tr>
<tr>
<td>B5</td>
<td>Building graffiti</td>
</tr>
<tr>
<td>B6</td>
<td>Structural integrity</td>
</tr>
<tr>
<td>B7</td>
<td>Broken roof</td>
</tr>
<tr>
<td>B8</td>
<td>Other</td>
</tr>
</tbody>
</table>

Environment/ Infrastructure

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Damaged sidewalk</td>
</tr>
<tr>
<td>E2</td>
<td>Damaged roads</td>
</tr>
<tr>
<td>E3</td>
<td>Overgrown vegetation</td>
</tr>
<tr>
<td>E4</td>
<td>Litter</td>
</tr>
<tr>
<td>E5</td>
<td>Illegal dumping</td>
</tr>
<tr>
<td>E6</td>
<td>Unkempt vacant areas</td>
</tr>
<tr>
<td>E7</td>
<td>Illegal parking</td>
</tr>
<tr>
<td>E8</td>
<td>Abandoned vehicle</td>
</tr>
<tr>
<td>E9</td>
<td>Graffiti (environment)</td>
</tr>
<tr>
<td>E10</td>
<td>Other</td>
</tr>
</tbody>
</table>
Some numbers...

- 1717 physical blight locations
- 69% environmental/infrastructural blight
- 31% property blight

Spatial video:
- 384 km
- 14 hrs 36 min
- 8 days
- Ø 26 kph

Geo-narratives:
- 46 participants (students/non-students)
- 25-30 min drive
Counts

- Abandoned building: 138
- Broken window/door: 40
- Blocked door/window: 265
- No: 54
- Building graffiti: 40
- Structural integrity: 282
- Building overgrowth: 48
- Other: 13

Overgrown vegetation: 193
Litter: 798
Dumping: 234
Unkempt areas: 186
Illegal parking: 16
Abandoned vehicle: 25
Infrastructural graffiti: 28
Other: 18

Environmental/Infrastructural blight
Density maps
Density maps

Normalized KDE - Blight collection with spatial video

Normalized KDE - 311 reported blight data

Blight collection with spatial video
- High: 100
- Low: 0

311 blight data
- High: 100
- Low: 0
Density maps

a. Abandoned properties
b. Blocked windows
c. Structural integrity
d. Dumping
e. Litter
f. Overgrown vegetation
Density maps
Density maps
Some insights

Overlay Blight and Crime
Fairfields Neighborhood

“Her grandson was shot and killed in a 2015 triple homicide outside the B’s Seafood convenience store about a block from her house — the same place her nephew was gunned down less than two years later.”
LSU study shows link between blighted property and homicide in Baton Rouge

By Donnie Lelake | February 11, 2019 at 10:44 AM CST - Updated February 11 at 10:44 AM

BATON ROUGE, LA (WAFB) - A new LSU study shows a link between homicide, blighted property and convenience stores in Baton Rouge.

The study began as a group project in Valaski's crime mapping class. Stephen Martinez, Valaski's student and co-author of the study, was interested in searching for data on whether or not murders were clustered near certain types of buildings. The project looked at homicides in Baton Rouge occurring in 2016.
Results

Urban blight → Safety perception
Survey

Places where test participants felt less safe

Categories influencing test participants’ crime perception from 1 (not at all) to 5 (most)
Geonarratives route

Commentary and feelings during geonarrative drive - student

These neighborhoods still look undeveloped, abandoned houses.
Here it looks dirty and abandoned.
There’s a person with their dog so I feel safe around.
I will try to avoid as much as I can these areas.
It’s easier for thieves to hide on trees waiting to assault you.
It’s a high school that can be a target for young teenagers walking after school.
These neighborhoods still look undeveloped, abandoned houses.
Here it looks dirty and abandoned.
There’s a person with their dog so I feel safe around.
I will try to avoid as much as I can these areas.
It’s easier for thieves to hide on trees waiting to assault you.
It’s a high school that can be a target for young teenagers walking after school.

Place
- Louisiana is one of the poorest states and can influence on crime rates

FEELING
- positive
- neutral
- negative

Route

Route for geo-narrative data collection

Mid City

Farfields
Sentiment analysis

Positive polarity for geonarratives route

Negative polarity for geonarratives route

KDE - positive
- Very low
- Low
- Moderate
- High
- Very high

Geonarratives route

KDE - negative
- Very low
- Low
- Moderate
- High
- Very high

Geonarratives route
Moments of Stress

Algorithm by Kyriakou, Resch et al. 2019
Relevance

- New geospatial technology as a methodology to improve the identification of crime-related variables and to explore urban safety;
- Identify physical urban blight indicators on a micro-scale;
- Collect contextual information in a standardized way and in a format that can be archived, so that they can be used in long term and comparative studies;
- Security improvements and enhancement of quality of life in Baton Rouge.
Future work

- Machine learning algorithms for image recognition;
- Automation of transcripts;
- Integration of UAV’s to record multiple facades of the property;
- Crime prediction models by including newly extracted information;
- Social media text analysis based on crime perception;
Let’s imagine
Join this event! Free registration!

The 2020 BARI Conference, Reimagined

Opening Keynote: The Other Impacts of Corona
Week 2: Co-Creation: Designing Together for Better Outcomes
Week 3: Collaboratively Building Climate Resilience
Week 4: Supporting the Vulnerable Few
Week 5 Keynote: Boston Innovation: Past Present and Future with Paul Grogan and Friends
Week 6: Strengthening the Commonwealth through Cross-Municipal Collaboration
Week 7: Making Housing in Greater Boston Work for Everyone
Week 8: Supporting Greater Boston’s Youth