Predicting Crime with Routine Activity Patterns Inferred from Social Media

Mohammad Al Boni and Matthew S. Gerber
Department of Systems and Information Engineering
University of Virginia, USA
SMC – Budapest
October, 10th, 2016
Executive Summary

Why listen for the next 15 minutes?
Executive Summary

Why listen for the next 15 minutes?

• The proposed models improved the prediction accuracy on 15 out of 20 crime types in Chicago, Illinois, USA.
Executive Summary

Why listen for the next 15 minutes?

• The proposed models improved the prediction accuracy on 15 out of 20 crime types in Chicago, Illinois, USA.

• We built crime prediction models using features:
  • Historical crime records.
  • Temporal features.
  • Micro-level movement patterns.
Executive Summary

Why listen for the next 15 minutes?

• The proposed models improved the prediction accuracy on 15 out of 20 crime types in Chicago, Illinois, USA.
• We built crime prediction models using features:
  • Historical crime records.
  • Temporal features.
  • Micro-level movement patterns.
• We estimated the movement patterns using Twitter and FourSquare data sources.
Big Picture: This is a Resource Allocation Problem
Big Picture: This is a Resource Allocation Problem
Big Picture: This is a Resource Allocation Problem

Chicago, Illinois, USA  
Theft Incidents 09/25/2016  
Crime Prediction (Hotspots)
Background and Motivations

• Late 60s, Cohen and Felson’s theory on routine activities [1]:
  • Motivated offender.
  • Suitable target.
  • Absence of capable guardian.

Background and Motivations

• Late 60s, Cohen and Felson’s theory on routine activities [1]:
  • Motivated offender.
  • Suitable target.
  • Absence of capable guardian.
• Geotagged social media implicitly describe individuals’ movements.

Background and Motivations

• Late 60s, Cohen and Felson’s theory on routine activities [1]:
  • Motivated offender.
  • Suitable target.
  • Absence of capable guardian.

• Geotagged social media implicitly describe individuals’ movements.

• **Objective:** We would like to improve upon existing crime prediction models by incorporating daily movement patterns of individuals on social media.

Background and Motivations

• Late 60s, Cohen and Felson’s theory on routine activities [1]:
  • Motivated offender.
  • Suitable target.
  • Absence of capable guardian.

• Geotagged social media implicitly describe individuals’ movements.

• **Objective:** We would like to improve upon existing crime prediction models by incorporating daily movement patterns of individuals on social media.

• We hypothesize that these activities (absent in previous research) are key contributors to crime outcomes.

Building crime prediction models using binary classifiers

Crime Prediction Model:

• Discretization
Building crime prediction models using binary classifiers

Crime Prediction Model:

- Discretization
Building crime prediction models using binary classifiers

Crime Prediction Model:

- Discretization
- Assign labels: NONE, T
Building crime prediction models using binary classifiers

Crime Prediction Model:

• Discretization

• Assign labels: NONE, T

• We obtained 197K incidents of 20 different crime types from the City of Chicago data portal between 2013-07-28 and 2014-04-14.
Building crime prediction models using binary classifiers

Crime Prediction Model:

• Discretization
• Assign labels: NONE, T
• We obtained 197K incidents of 20 different crime types from the City of Chicago data portal between 2013-07-28 and 2014-04-14.
• Binary classifier.

\[
Pr(Label_p = T|f_1(\theta_1), \ldots, f_n(\theta_n)) = F(f_1(\theta_1), \ldots, f_n(\theta_n))
\]

\[
F(f_1(\theta_1), \ldots, f_n(\theta_n)) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i f_i(\theta_i))}}
\]
Building crime prediction models using binary classifiers

**Crime Prediction Model:**

- Discretization
- Assign labels: NONE, T
- We obtained 197K incidents of 20 different crime types from the City of Chicago data portal between 2013-07-28 and 2014-04-14.
- Binary classifier.
- Correlating movement patterns to Crime.

![Diagram showing movement patterns and crime incidents](image)

- Point 1: 8:30 PM, Home
- Point 2: 9:00 PM, Bar
- Point 3: 10:00 PM, Bar
- Point 4: 11:30 PM, Bar
- Point 5: 11:30 PM, Theft
- Point 6: 12:30 AM, Home
Estimating activity patterns from Tweets

Sequence of Tweets:
We collected >9M GPS-tagged tweets authored by >200K different users within Chicago between 2013-07 and 2014-04.
Estimating activity patterns from Tweets

Sequence of Tweets:

1- Split tweets using activity diagrams:
Estimating activity patterns from Tweets

Sequence of Tweets:
1- Split tweets using activity diagrams:
Estimating activity patterns from Tweets

Sequence of Tweets:
1- Split tweets using activity diagrams:
Estimating activity patterns from Tweets

Sequence of Tweets:
1- Split tweets using activity diagrams:
2- Map to venues within 5-meters:
   We extracted >200K different logical places for Chicago using FourSquare.
Estimating activity patterns from Tweets

Sequence of Tweets:

1- Split tweets using activity diagrams:

2- Map to venues within 5-meters:
   Home → The Common Cup → CTA Bus Stop 1006 → CTA Bus Stop 1120 → Lincoln Building

3- Get venue categories:
   Each FourSquare place has a category. FourSquare provides a hierarchical list of categories with 10 at the top:

<table>
<thead>
<tr>
<th></th>
<th>Residence</th>
<th>Professional &amp; Other Places</th>
<th>Shop &amp; Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel &amp; Transport</td>
<td>Nightlife Spot</td>
<td>Outdoors &amp; Recreation</td>
<td></td>
</tr>
<tr>
<td>Arts &amp; Entertainment</td>
<td>College &amp; University</td>
<td>Food</td>
<td>Event</td>
</tr>
</tbody>
</table>
Estimating activity patterns from Tweets

Sequence of Tweets:

1- Split tweets using activity diagrams:

2- Map to venues within 5-meters:
   Home → The Common Cup → CTA Bus Stop 1006 → CTA Bus Stop 1120 → Lincoln Building

3- Get venue categories:
   Residence → Food → Travel & Transport → Travel & Transport → Professional & Other Places

4- Represent activities using bag-of-words and compute Venue Frequency-Inverse Route Frequency features:
   - We use BOW representation: each venue type = dimension.
   - We use VF-IRF weighting scheme:
     - VF: a venue type is more important if it occurs more frequently in a route.
       - $VF(v, r) = c(v, r)$
     - IRF: a venue type is more discriminative if it occurs only in fewer routes.
       - $IRF(t) = 1 + \log\left(\frac{N}{rf(v)}\right)$

   Total number of routes in training
   Number of routes containing venue $v$
Estimating activity patterns from Tweets

**Sequence of Tweets:**

1- Split tweets using activity diagrams:

2- Map to venues within 5-meters:
   Home → The Common Cup → CTA Bus Stop 1006 → CTA Bus Stop 1120 → Lincoln Building

3- Get venue categories:
   Residence → Food → Travel & Transport → Travel & Transport → Professional & Other Places

4- Represent activities using bag-of-words and compute Venue Frequency-Inverse Route Frequency features:
   Compute Venue Frequency: \(VF = \{1, 1, 0, 2, 1, 0, 0, 0, 0\}\)
Estimating activity patterns from Tweets

Sequence of Tweets:

1- Split tweets using activity diagrams:
2- Map to venues within 5-meters:
   Home → The Common Cup → CTA Bus Stop 1006 → CTA Bus Stop 1120 → Lincoln Building
3- Get venue categories:
   Residence → Food → Travel & Transport → Travel & Transport → Professional & Other Places
4- Represent activities using bag-of-words and compute Venue Frequency-Inverse Route Frequency features:

Compute Venue Frequency: \( VF = \{1, 1, 0, 2, 1, 0, 0, 0, 0, 0\} \)

Assume: IRF: Residence= 1.84, Food= 1.64, Travel & Transport= 1.7, and Professional & Other Places= 1.57

Compute final VF-IRF vector: \( VF-IRF = \{1.84, 1.57, 0, 3.4, 1.64, 0, 0, 0, 0, 0\} \)
Building crime prediction models by incorporating the daily activities

Crime Prediction Model:

\[ Pr(Label_p = T|f_1(\theta_1), ..., f_{14}(\theta_{14})) = F(f_1(\theta_1), ..., f_{14}(\theta_{14})) \]

\[ F(f_1(\theta_1), ..., f_{14}(\theta_{14})) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{14} \beta_i f_i(\theta_i))}} \]

- \( f_1(p) \): Kernel Density Estimation (KDE) – Historical crime records (Gerber 2014).
Building crime prediction models by incorporating the daily activities

**Crime Prediction Model:**

\[
Pr(Label_p = T|f_1(\theta_1), ..., f_{14}(\theta_{14})) = F(f_1(\theta_1), ..., f_{14}(\theta_{14}))
\]

\[
F(f_1(\theta_1), ..., f_{14}(\theta_{14})) = \frac{1}{1 + e^{- (\beta_0 + \sum_{i=1}^{14} \beta_i f_i(\theta_i))}}
\]

- \(f_1(\{p\})\): Kernel Density Estimation (KDE) – Historical crime records (Gerber 2014).
- \(f_2(\{t_c\})\) and \(f_2(\{t_c\})\): two features to capture the temporal component of the prediction.
Building crime prediction models by incorporating the daily activities

**Crime Prediction Model:**

\[
Pr(\text{Label}_p = T|f_1(\theta_1), ..., f_{14}(\theta_{14})) = F(f_1(\theta_1), ..., f_{14}(\theta_{14}))
\]

\[
F(f_1(\theta_1), ..., f_{14}(\theta_{14})) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{14} \beta_i f_i(\theta_i))}}
\]

- \(f_1(\{p\})\): Kernel Density Estimation (KDE) – Historical crime records (Gerber 2014).
- \(f_2(\{t_c\})\) and \(f_2(\{t_c\})\): two features to capture the temporal component of the prediction.
- \(f_3(\{p, t_1, t_2\}), ..., f_{13}(\{p, t_1, t_2\})\): accumulative TF-IDF features of routine activities:

\[
f_{i+3}(\{p, t_1, t_2\}) = \sum_{r \in R(p, t_1, t_2)} VF - IRF(c_i, r)
\]
Building crime prediction models by incorporating the daily activities

Crime Prediction Model:

\[
Pr(\text{Label}_p = T| f_1(\theta_1), \ldots, f_{14}(\theta_{14})) = F(f_1(\theta_1), \ldots, f_{14}(\theta_{14}))
\]

\[
F(f_1(\theta_1), \ldots, f_{14}(\theta_{14})) = \frac{1}{1 + e^{-\left(\beta_0 + \sum_{i=1}^{14} \beta_i f_i(\theta_i)\right)}}
\]

- \(f_1(\{p\})\): Kernel Density Estimation (KDE) – Historical crime records (Gerber 2014).
- \(f_2(\{t_c\})\) and \(f_2(\{t_c\})\): two features to capture the temporal component of the prediction.
- \(f_3(\{p, t_1, t_2\}), \ldots, f_{13}(\{p, t_1, t_2\})\): accumulative TF-IDF features of routine activities:
  \[
  f_{i+3}(\{p, t_1, t_2\}) = \sum_{r \in R(p, t_1, t_2)} VF - IRF(c_i, r)
  \]
- \(f_{14}(\{p, t_1, t_2\}) = \sum_{r \in R(p, t_1, t_2)} |r|\)
Evaluating Crime Prediction Models Using Surveillance Plots

# of observed future incidents = 40
Evaluating Crime Prediction Models Using Surveillance Plots

# of observed future incidents = 40

Top 10% most-threatened area according to the prediction.
Evaluating Crime Prediction Models Using Surveillance Plots

# of observed future incidents = 40

# of incidents captured by hot spots = 8

20% of the future incidents

Top 10% most-threatened area according to the prediction.
Evaluating Crime Prediction Models Using Surveillance Plots

# of observed future incidents = 40

# of incidents captured by hot spots = 26

65% of the future incidents

Top 40% most-threatened area according to the prediction.
Evaluating Crime Prediction Models Using Surveillance Plots

Evaluating Crime Prediction Models Using Surveillance Plots


Overall, Model 2 has better performance.
Evaluating Crime Prediction Models Using Surveillance Plots

Within the top 20% most threatened area, Model 1 is better.
Improved prediction performance on 15 out of 20 crime types

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>KDE+T (AUC)</th>
<th>+R (AUC)</th>
<th>-Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEAPONS VIOLATION</td>
<td>0.65</td>
<td>0.68</td>
<td>0.09@40</td>
</tr>
<tr>
<td>PUBLIC PEACE VIOLATION</td>
<td>0.65</td>
<td>0.67</td>
<td>0.09@25</td>
</tr>
</tbody>
</table>
Improved prediction performance on 15 out of 20 crime types

- 9 crime types with peak gains within half of the surveillance area.

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>KDE+T (AUC)</th>
<th>+R (AUC)</th>
<th>-Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>THEFT</td>
<td>0.65</td>
<td>0.68</td>
<td>0.06@54</td>
</tr>
<tr>
<td>BATTERY</td>
<td>0.71</td>
<td>0.74</td>
<td>0.04@43</td>
</tr>
<tr>
<td>CRIMINAL DAMAGE</td>
<td>0.65</td>
<td>0.67</td>
<td>0.03@64</td>
</tr>
<tr>
<td>OTHER OFFENSE</td>
<td>0.64</td>
<td>0.69</td>
<td>0.08@50</td>
</tr>
<tr>
<td>BURGLARY</td>
<td>0.68</td>
<td>0.7</td>
<td>0.07@51</td>
</tr>
<tr>
<td>ASSAULT</td>
<td>0.67</td>
<td>0.71</td>
<td>0.07@51</td>
</tr>
<tr>
<td>MOTOR VEHICLE THEFT</td>
<td>0.65</td>
<td>0.68</td>
<td>0.06@54</td>
</tr>
<tr>
<td>ROBBERY</td>
<td>0.69</td>
<td>0.73</td>
<td>0.07@54</td>
</tr>
<tr>
<td>OFFENSE INVOLVING CHILDREN</td>
<td>0.58</td>
<td>0.58</td>
<td>0.04@50</td>
</tr>
</tbody>
</table>
Improved prediction performance on 15 out of 20 crime types

- 4 crime types with peak gains within quarter of the surveillance area.

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>KDE+T (AUC)</th>
<th>+R (AUC)</th>
<th>-Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>NARCOTICS</td>
<td>0.76</td>
<td>0.83</td>
<td>0.18@25</td>
</tr>
<tr>
<td>DECEPTIVE PRACTICE</td>
<td>0.69</td>
<td>0.75</td>
<td>0.16@25</td>
</tr>
<tr>
<td>CRIMINAL TRESPASS</td>
<td>0.65</td>
<td>0.72</td>
<td>0.14@25</td>
</tr>
<tr>
<td>PROSTITUTION</td>
<td>0.66</td>
<td>0.68</td>
<td>0.13@25</td>
</tr>
</tbody>
</table>
Outdoors & recreation, nightlife spot, and residence venues have positive correlation with prostitution crime type

**Coefficient Analysis:**

<table>
<thead>
<tr>
<th>Location</th>
<th>Coefficient</th>
<th>Location</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence</td>
<td>6.92</td>
<td>Nightlife Spot</td>
<td>7.54</td>
</tr>
<tr>
<td>Professional &amp; Other Places</td>
<td>-5.61</td>
<td>Outdoors &amp; Recreation</td>
<td>8.87</td>
</tr>
<tr>
<td>Shop &amp; Service</td>
<td>-3.53</td>
<td>Arts &amp; Entertainment</td>
<td>-3.19</td>
</tr>
<tr>
<td>Travel &amp; Transport</td>
<td>1.43</td>
<td>College &amp; University</td>
<td>1.89</td>
</tr>
<tr>
<td>Food</td>
<td>-3.78</td>
<td>Event</td>
<td>4.79</td>
</tr>
</tbody>
</table>
College & University venues have positive correlation with prostitution, weapons violation, and sex offense crime types

Coefficient Analysis:

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>Coefficient</th>
<th>Crime Type</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>NARCOTICS</td>
<td>-4.09</td>
<td>GAMBLING</td>
<td>-15.16</td>
</tr>
<tr>
<td>DECEPTIVE PRACTICE</td>
<td>-3.57</td>
<td>MOTOR VEHICLE THEFT</td>
<td>-6.76</td>
</tr>
<tr>
<td>CRIMINAL TRESPASS</td>
<td>-1.85</td>
<td>THEFT</td>
<td>-3.64</td>
</tr>
<tr>
<td>PROSTITUTION</td>
<td>1.89</td>
<td>OFFENSE INVOLVING CHILDREN</td>
<td>0.12</td>
</tr>
<tr>
<td>PUBLIC PEACE VIOLATION</td>
<td>1.56</td>
<td>INTERFERENCE WITH PUBLIC OFFICER</td>
<td>-1.66</td>
</tr>
<tr>
<td>WEAPONS VIOLATION</td>
<td>0.79</td>
<td>BATTERY</td>
<td>-5.11</td>
</tr>
<tr>
<td>OTHER OFFENSE</td>
<td>-4.54</td>
<td>CRIMINAL DAMAGE</td>
<td>-3.74</td>
</tr>
<tr>
<td>ASSAULT</td>
<td>-3.52</td>
<td>LIQUOR LAW VIOLATION</td>
<td>-12.5</td>
</tr>
<tr>
<td>ROBBERY</td>
<td>-3.09</td>
<td>SEX OFFENSE</td>
<td>0.51</td>
</tr>
<tr>
<td>BURGLARY</td>
<td>-4.89</td>
<td>HOMICIDE</td>
<td>-14.12</td>
</tr>
</tbody>
</table>
Summary & Future Work

• The proposed models improved the prediction accuracy on 15 out of 20 crime types in Chicago, Illinois, USA.

• We built crime prediction models using features:
  • Historical crime records.
  • Temporal features.
  • Micro-level movement patterns.

• We estimated the movement patterns using Twitter and FourSquare data sources.

• Future work:
  • Include higher N-gram features (i.e., to capture bar->bar->bar).
  • Use more detailed venue types.
Temporal Features