Fostering Smart City Development in Developing Nations: A Crime Series Data Analytics Approach

Omowunmi Isafiade
University of the Western Cape
oisafiade@uwc.ac.za
Outline

• Introduction
  – Motivation and challenges
• Crime Control in South Africa
  – Current practice and gaps
• Proposed Intervention: CriClust System
  – Model formulation and design methodology
  – Results and discussion
• Conclusion and Outlook
• References
• Acknowledgements
Rapid urbanisation: More people live in cities than rural areas

Fig: Growth of African cities
• More than 50% of the world’s population lives in cities.
• Continued rural-urban migration forecasted up to year 2025.

Fig: Trend in South Africa
Smart City: Meeting the challenges of rapid urbanisation

- Increase in crime anticipated with rapid urbanisation.
- Deterring crime is a top priority for realising a sustainable “safe and smart” city.
- The use of armed weapons is not sufficient to tackle crime.

**Smart city**: using urban informatics and technology to improve the quality and efficiency of urban operation and services.
• Fig: Distribution of crime across provinces in South Africa
### Current Practice

- Random patrols at locations
- Manual means of data capture & processing (using excel software)
- Accumulated data is transferred to provincial level for processing
- CrimeHub statistics [Institute for Security Studies (ISS)]

### Gaps/Limitations

- **Police: Citizen** ratio is 1:347 (288 police per 100,000 citizen)
- Limited technological tool for pattern detection
- Delay in knowledge discovery (inaccuracies)
- Mitigation practices hindered due to lack of domain experts and technological tools (e.g., Analyst's Notebook)
- General background information – may not be actionable
Problem Statement: Challenges of squeezing crime to zero

- Despite the vast resources allocated to crime, people still fall victim of crime
- Plethora of under-utilised crime reports archived by public safety.
- Manual means adopted at local stations is a huge constraint to effective policing in developing nations (e.g. South Africa).
- Need to promote knowledge-driven decision support for public safety improvement in developing nations.
- Crime series pattern (CSP) detection is less explored in developing nations
CriClust: Crime Series Pattern (CSP) detection

- Depiction of serial predator in related crime scenarios in a city.

- Research shows that many crimes are due to repeat (serial) offenders: crime series.

- Crime series are crimes committed by same offender.

- If patterns are identified timeously police can prevent further recurrence.

- Several tools exist but mostly able to estimate background information.
Crime Series Pattern (CSP) Detection

- CriClust serves to assist in CSP detection using rape data.
- However, can extend to other forms of crime
- Issues around rape and sexual violence still an ongoing concern in South African communities.
- Hence, crucial to devise smart means of assisting police in developing nations
Phases in CriClust System

Fig: An overview of research phases in CriClust System
CriClust: Problem definition

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>...</th>
<th>$A_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_i$</td>
<td>$A_{1_i}$</td>
<td>$A_{2_i}$</td>
<td>...</td>
<td>$A_{n_i}$</td>
</tr>
<tr>
<td>$C_j$</td>
<td>$A_{1_j}$</td>
<td>$A_{2_j}$</td>
<td>...</td>
<td>$A_{n_j}$</td>
</tr>
<tr>
<td>$C_k$</td>
<td>$A_{1_k}$</td>
<td>$A_{2_k}$</td>
<td>...</td>
<td>$A_{n_k}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- Let $C$ be a set of crime objects, where each $i$ in $C$ is defined by a set of attributes $A$, our interest lies in crime objects that exhibit a coherent pattern on a subset of $A$. 
Crime Series Pattern Detection: Algorithmic process

- $\alpha = \frac{360^0 \text{ (in a circle)}}{7 \text{ (days in a week)}}$
- $\alpha$ is the angle between each pair of days
- The 2-D component is relevant because a 1-D component will assume that Sunday is far from Monday.
CriClust: Learning the similarity graph

Fig: Flow of highly connected Subgraphs (HCS)

Fig: A depiction of crime cluster detected by HCS
Overview: CriClust System Visualisation

CriClust | Crime Data Mining

admin15

Login
CriClust: Scalability and trend of series observed

Fig: Scalability trend

Fig: Trend of series observed across locations
Characterising features emerging for each series

<table>
<thead>
<tr>
<th>S/N</th>
<th>Location</th>
<th>PDE(%)</th>
<th>Day</th>
<th>Time</th>
<th>Vic</th>
<th>Sus</th>
<th>VAge</th>
<th>SAge</th>
<th>SFr</th>
<th>Mot</th>
<th>MO</th>
<th>HCol</th>
<th>Mask</th>
<th>Sub-Ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mowbray</td>
<td>35 (S1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>65 (S2)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>CapeTown</td>
<td>50 (S1)</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Central</td>
<td>50 (S2)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Wynberg</td>
<td>40(S1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>34(S2)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>26(S3)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Grassy-</td>
<td>21(S1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Park</td>
<td>79(S2)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
# Systematic comparison of CriClust with existing research

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exploratory basis</td>
<td>Crime linkage</td>
<td>Crime series detection</td>
<td>Serial criminal pattern detection</td>
<td>Crime linkage</td>
<td>Crime series detection</td>
</tr>
<tr>
<td>2</td>
<td>Crime explored</td>
<td>Breaking &amp; entering crimes</td>
<td>Burglary (housebreaking)</td>
<td>Armed robberies</td>
<td>Burglary crimes</td>
<td>Sexual crime</td>
</tr>
<tr>
<td>3</td>
<td>Modelling approach</td>
<td>Statistical approach</td>
<td>Conventional optimisation</td>
<td>Neural Network (NN)</td>
<td>Bayesian Network (BN)</td>
<td>Dual threshold scheme &amp; graphical model</td>
</tr>
<tr>
<td>4</td>
<td>Techniques used</td>
<td>Bayes factor, Hierarchical clustering</td>
<td>Integer linear programming, clustering, BFS</td>
<td>Cascaded network of Kohonen NN</td>
<td>Bayes Network</td>
<td>Geometric projection, HCS clustering</td>
</tr>
<tr>
<td>5</td>
<td>Empirical observation</td>
<td>Posterior odds, Bayes factor &amp; number of clusters</td>
<td>Map location of series, pattern space, precision &amp; recall</td>
<td>Percentage of predicted &amp; actual patterns</td>
<td>Posterior probabilities &amp; BN</td>
<td>Map (PDE, PSE) of series, scalability, precision &amp; recall</td>
</tr>
</tbody>
</table>
Summary and Conclusion

- Challenge of crime is magnified in resource constraint settings.
- Police need to be empowered with context-aware and cost-effective technologies for effective policing.
- Crime series detection is less explored in developing nations.
- CriClust serves to assist in crime series identification, using a dual threshold mechanism and geometric projection.
- CriClust is not a panacea but can assist with underperformance in policing.
- CriClust is to be considered for deployment with the police, and there is an ongoing collaboration with an NGO on community policing.
References


References


Acknowledgements
Thank You