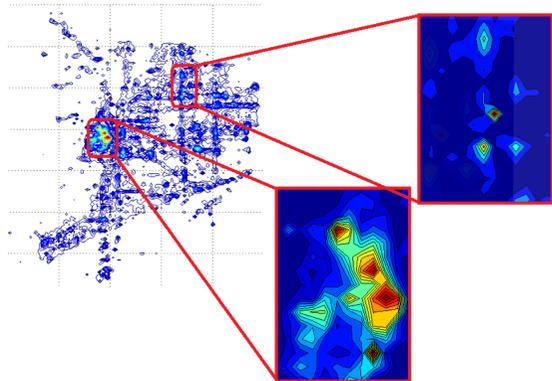


Ajitesh Srivastava, Viktor K. Prasanna
Data Science Lab (dslab.usc.edu)

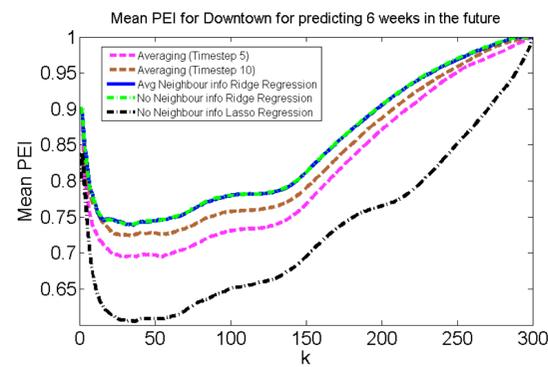
Scalable solutions for large-scale real-time crime prevention with theoretical guarantees

Portland Crime Prediction: Preliminary Analysis



Crime prediction for dense crime zones using

- Moving average
- Linear Regressions

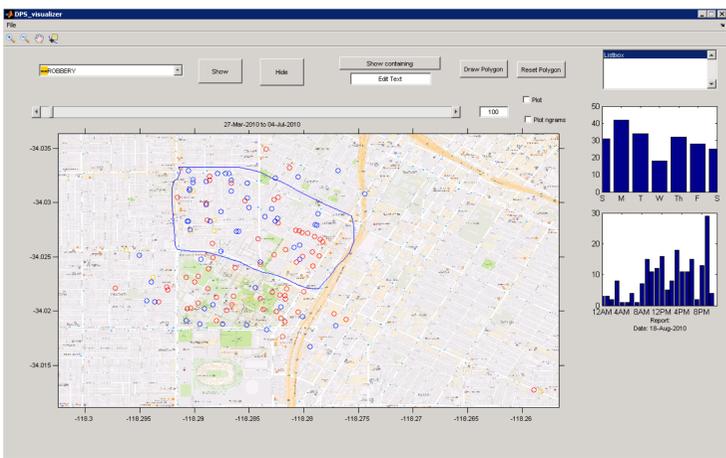


Lessons Learned

- Moving average is hard to beat for short term predictions
- Unable to leverage immediate neighborhood
- What is a good metric?

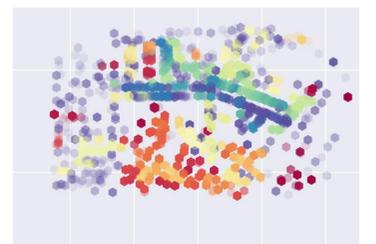
Crime Visualizer

- Spatio-temporal exploration of crime
- Crime Statistics
- Keyword Search



Improved Crime Prediction with a Novel Metric

Spatial clustering of crime to identify heterogeneous regions of crime

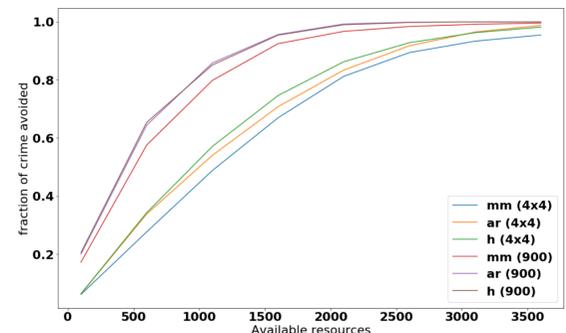


How to compare two methods?

- MSE, AUC, do not apply when regions are partitioned differently

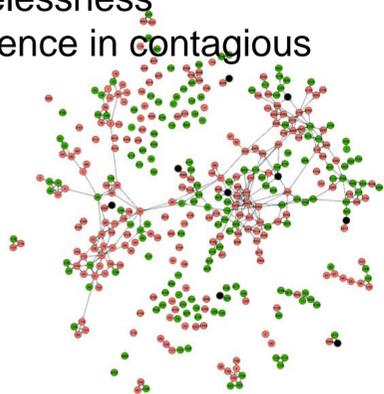
Resource allocation based metric

- Finite resources k : each prevents certain number of crimes in certain units of area
- Use predictions to optimally assign resources
- Use the ground truth to measure actual number of crimes prevented



Minimizing Violence among Homeless

- 56% of homeless youth in LA involved in violence in 2016
- Prevents youth from exiting homelessness
- Violence is contagious



Model

- At every time step a node u picks a node v and takes its state
- With probability θ : v is selected from neighborhood of u
 - With probability $1-\theta$: v is selected from outside of the neighborhood

Given: A graph $G(V, E)$ and a set of nodes S that are violent and Intervention resources - an integer k
To find: $T \subseteq S$ such that $|T| = k$, and turning the nodes in T into non-violent minimizes expected number of violent nodes

Greedy Algorithm guarantees the optimal solution

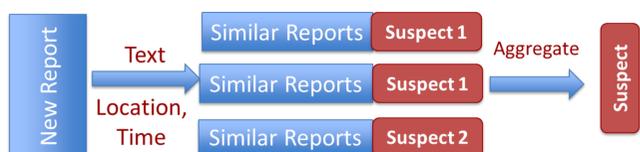
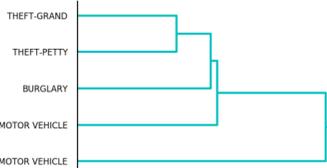
Table 1: Top 10 seeds for various values of θ output by Greedy Minimization

θ	Selected Seeds										$\mathbb{E}(I_V^T x_t)$
1	67	82	14	165	100	2222	156	61	90	2176	83.5783
0.9	67	82	14	61	165	2222	100	156	90	174	62.9762
0.8	67	82	61	14	2222	156	165	100	90	2110	51.3933
0.7	67	82	61	14	2222	156	2110	165	100	2138	45.7105
0.6	67	61	82	2222	156	14	2110	2138	165	174	44.3753
0.5	67	61	82	2110	2222	156	14	2138	2126	174	46.8745

Text Mining of Crime Reports

Identify similar crimes

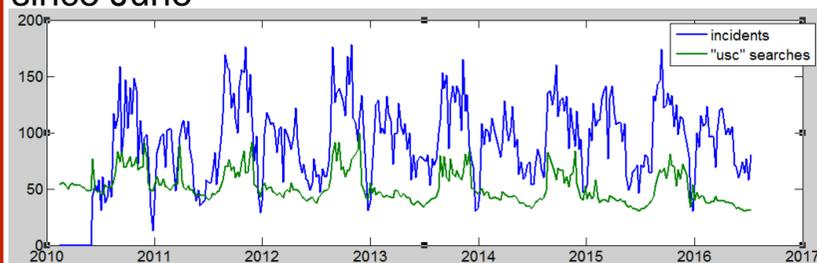
- Cluster to improve prediction
- To find suspects



Future Work

Effect of Social Media

- Ongoing collection of geo-tagged tweets in LA since June



Delivering Softwares

- Department of Public Safety, USC
- South Park Business District