Data Analytics

Predictive Policing

Visualization for Law and Order

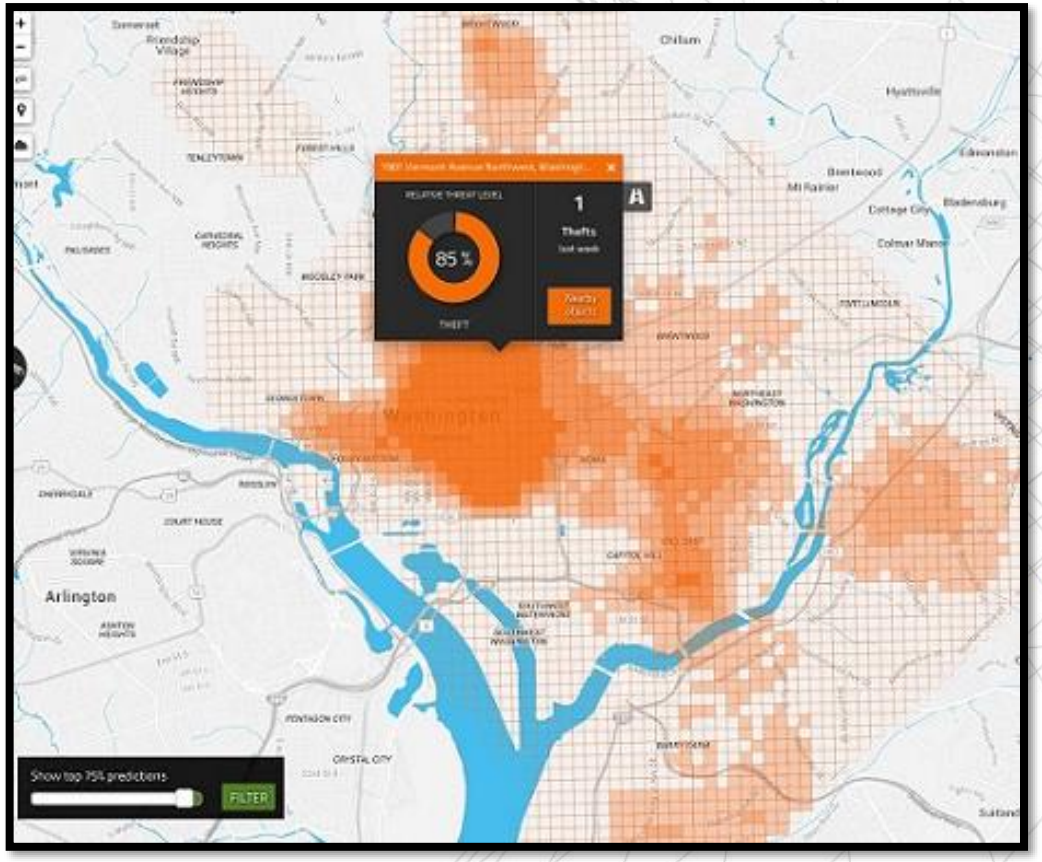
Guodong Lyu

NUS Business School

Note. This lecture is designed based on the text book “The Analytics Edge” by Allison K. O'Hair, Dimitris Bertsimas, William R. Pulleyblank

Introduction

- Predictive policing refers to the usage of mathematical, predictive and analytical techniques in law enforcement to identify potential criminal activity. Predictive policing methods fall into four general categories: methods for predicting crimes, methods for predicting offenders, methods for predicting perpetrators' identities, and methods for predicting victims of crime.



Introduction

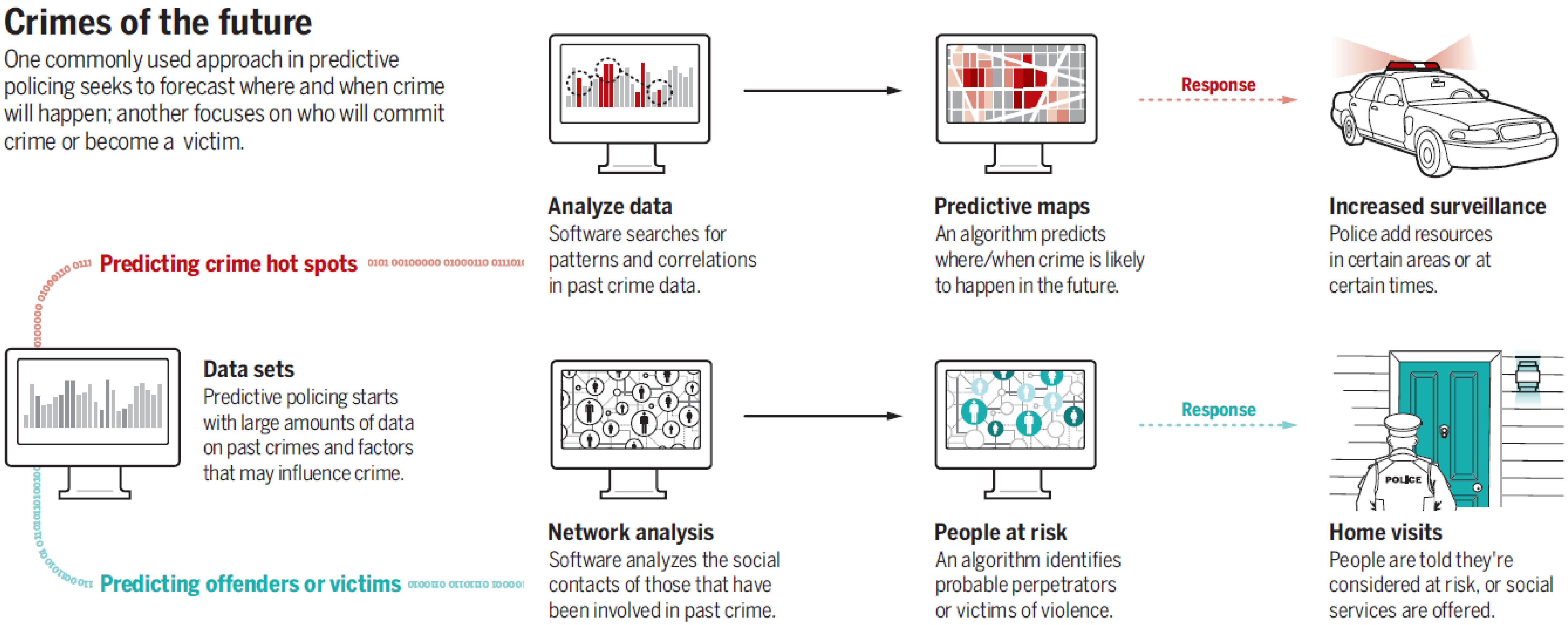
- In the past, human judgment and experience was the only tool in identifying patterns in criminal behaviour.

- The explosion of computerized data affects all parts of society, including law and order.

- Police forces around the US and the world are augmenting human judgment with analytics – sometimes described as “predictive policing”.

- In November 2011, TIME Magazine named predictive policing as one of the 50 best inventions of 2011.

Research Topics



- Visualization bridges the gap between the data and mathematics and the end user.

Predicting Crime Hot Spots

1. Dynamic Crime Index

- Regression based
- Aggregated level
- Historical average

2. Visualization Tools

- Line plot
- Heat map

3. Self-exciting point process model

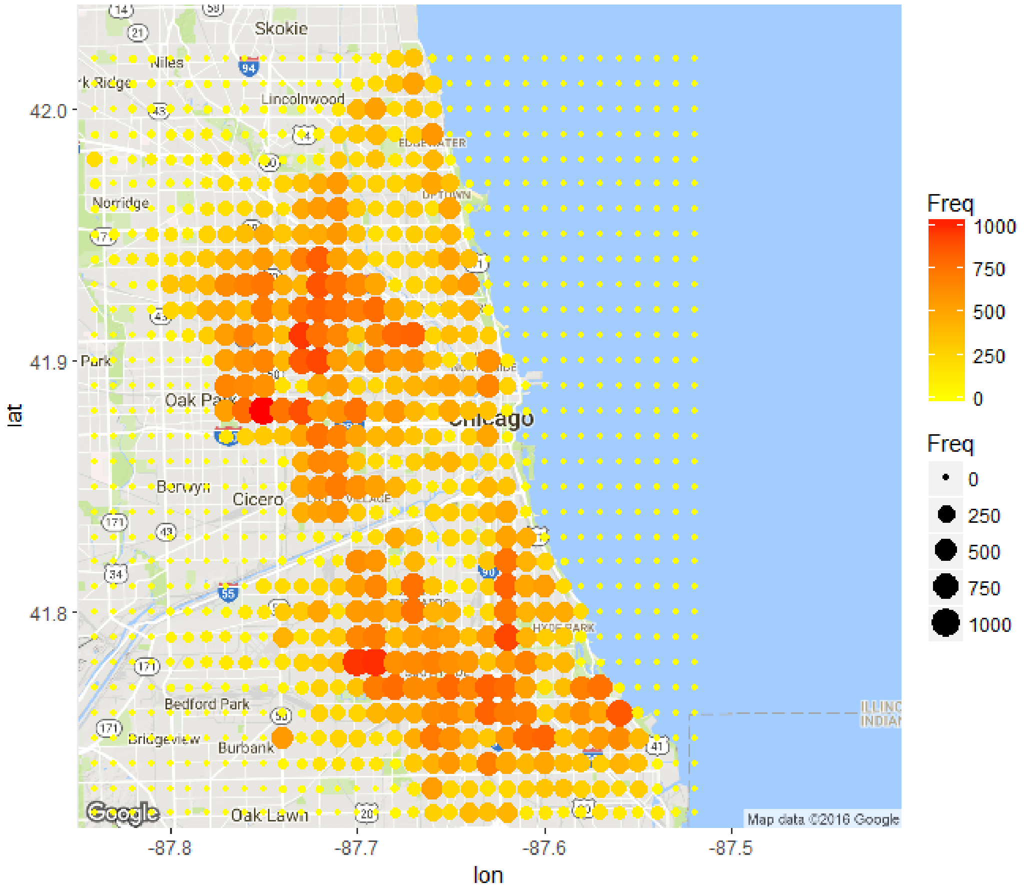
- Spatial-temporal patterns
- Earthquake (event prediction
- Page-rank approach

Predicting Crime Hot Spots

Dynamic Crime Index

Predicts the expected amount of crime for each neighborhood in the city/country.

- Crime group: homicide, assault, robbery, burglary, larceny and auto theft;
- Districts: location information, including longitude and latitude;
- Time: divides the data into training set and testing set.



Predicting Crime Hot Spots

Dynamic Crime Index: dataset

	Date	Latitude	Longitude	Weekday	Hour	Year	Month
1	2012-12-31 23:15:00	41.75628	-87.62164	Monday	23	2012	12
2	2012-12-31 22:00:00	41.89879	-87.66130	Monday	22	2012	12
3	2012-12-31 22:00:00	41.96919	-87.76767	Monday	22	2012	12
4	2012-12-31 22:00:00	41.76933	-87.65773	Monday	22	2012	12
5	2012-12-31 21:30:00	41.83757	-87.62176	Monday	21	2012	12
6	2012-12-31 20:30:00	41.92856	-87.75400	Monday	20	2012	12
7	2012-12-31 20:10:00	41.73206	-87.56481	Monday	20	2012	12
8	2012-12-31 20:00:00	41.79251	-87.61932	Monday	20	2012	12
9	2012-12-31 19:00:00	41.86362	-87.70910	Monday	19	2012	12
10	2012-12-31 18:00:00	41.89008	-87.65882	Monday	18	2012	12
11	2012-12-31 18:00:00	41.86542	-87.72025	Monday	18	2012	12
12	2012-12-31 17:00:00	41.76506	-87.66066	Monday	17	2012	12
13	2012-12-31 16:30:00	41.82054	-87.68490	Monday	16	2012	12
14	2012-12-31 16:20:00	41.78455	-87.68393	Monday	16	2012	12
15	2012-12-31 16:15:00	41.89349	-87.62160	Monday	16	2012	12
16	2012-12-31 16:00:00	41.74444	-87.68835	Monday	16	2012	12
17	2012-12-31 16:00:00	41.96477	-87.72165	Monday	16	2012	12
18	2012-12-31 16:00:00	41.88435	-87.63108	Monday	16	2012	12

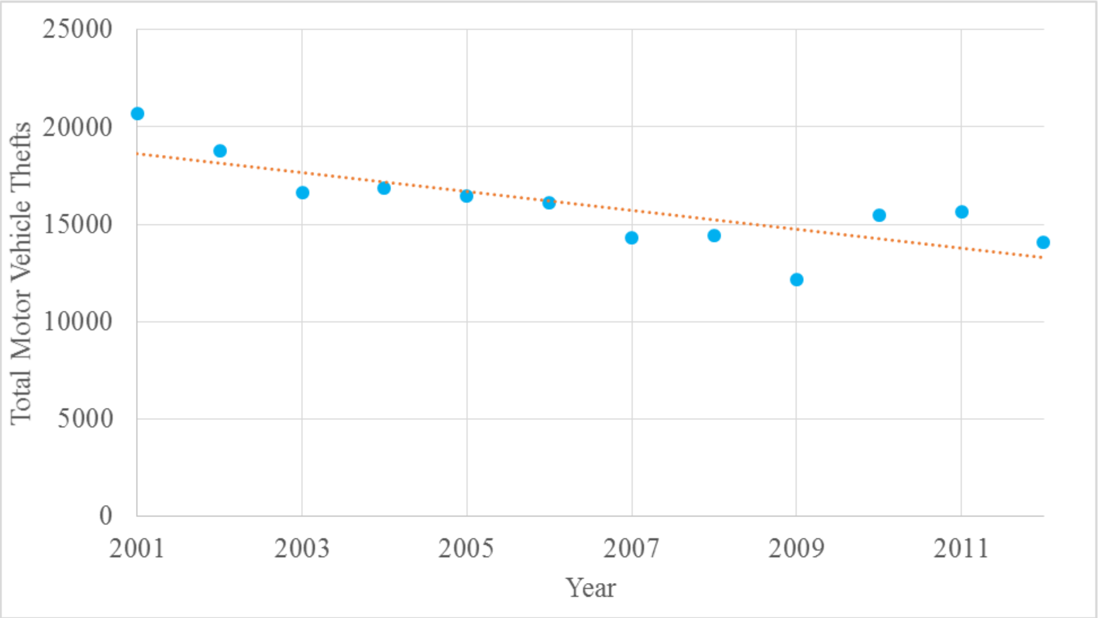
Data from Chicago about motor vehicle thefts

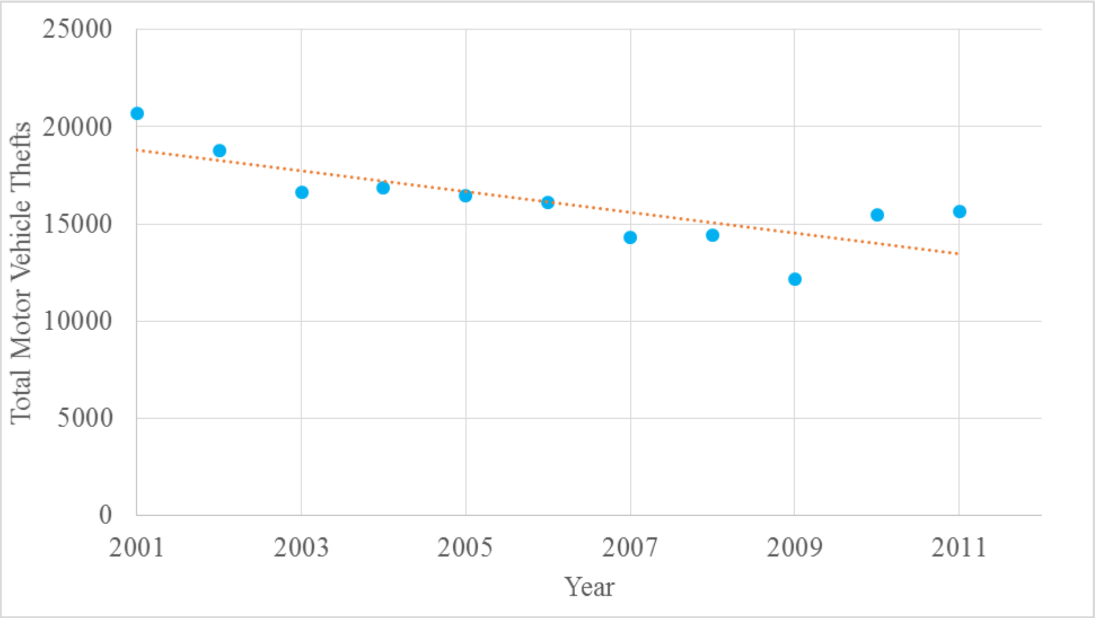
	State	Population	PopulationDensity	Murders	GunMurders	GunOwnership
1	Alabama	4779736	94.650	199	135	0.517
2	Alaska	710231	1.264	31	19	0.578
3	Arizona	6392017	57.050	352	232	0.311
4	Arkansas	2915918	56.430	130	93	0.553
5	California	37253956	244.200	1811	1257	0.213
6	Colorado	5029196	49.330	117	65	0.347
7	Connecticut	3574097	741.400	131	97	0.167
8	Delaware	897934	470.700	48	38	0.255
9	District of Columbia	601723	10298.000	131	99	0.036
10	Florida	19687653	360.200	987	669	0.245
11	Georgia	9920000	172.500	527	376	0.403
12	Hawaii	1360301	216.800	24	7	0.067
13	Idaho	1567582	19.500	21	12	0.553
14	Illinois	12830632	231.900	453	364	0.202
15	Indiana	6483802	182.500	198	142	0.391
16	Iowa	3046355	54.810	38	21	0.429
17	Kansas	2853118	35.090	100	63	0.421
18	Kentucky	4339367	110.000	180	116	0.477

Murders in the U.S.

Predicting Crime Hot Spots

Dynamic Crime Index: time-based pattern

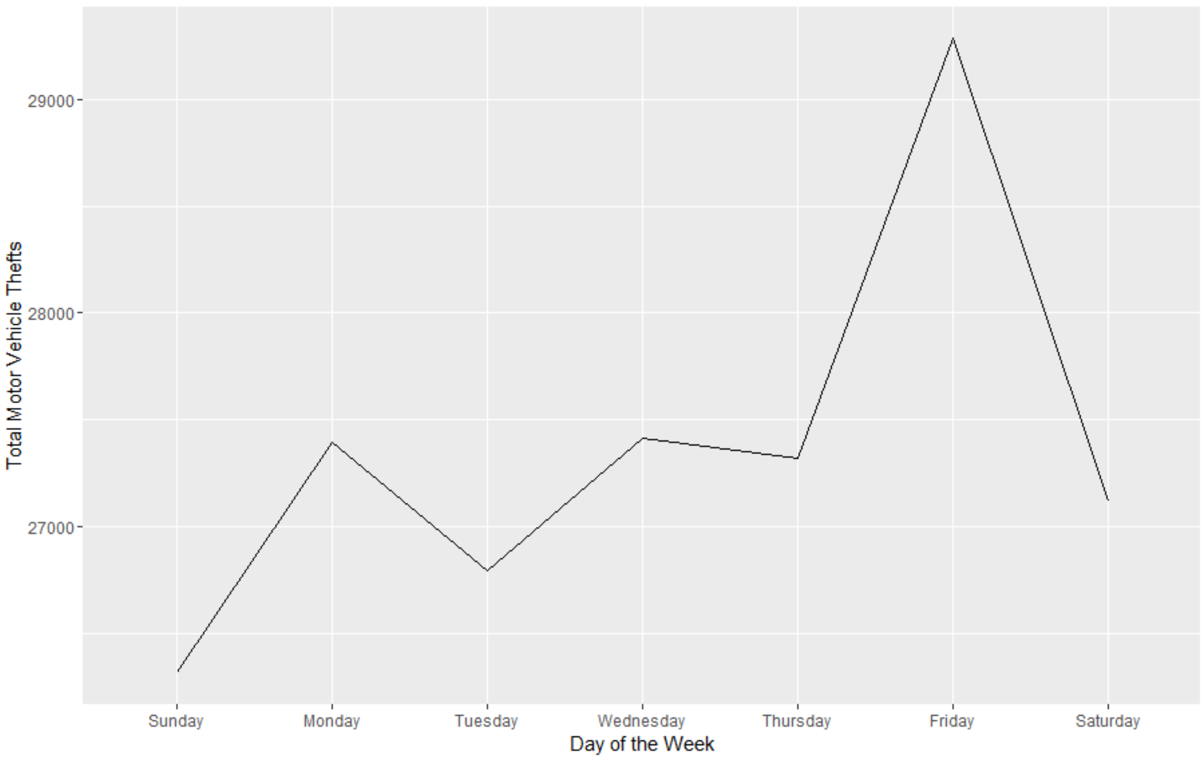


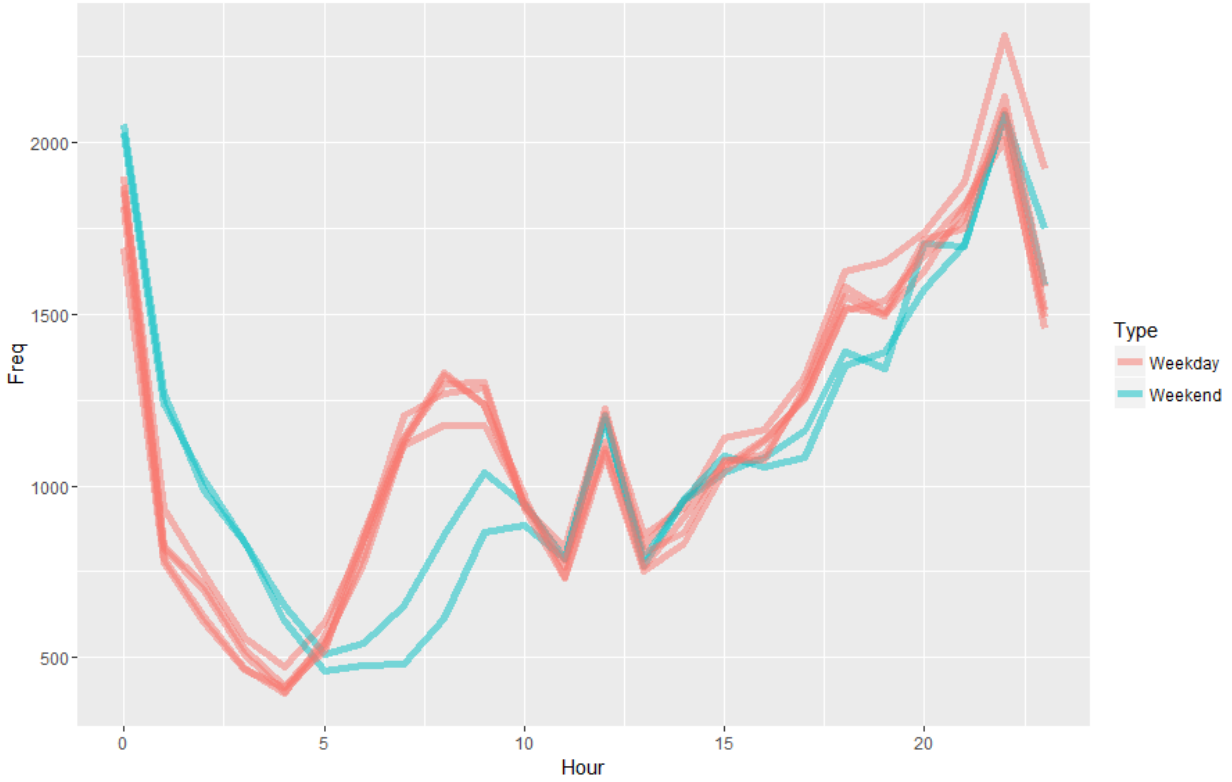


Prediction(2012) = 12939
Actual(2012) = 14092

Predicting Crime Hot Spots

Dynamic Crime Index: time-based pattern





Predicting Crime Hot Spots

Dynamic Crime Index: time-based pattern

- We can replace our **x-axis** with the **hour of the day**, and have a different **line for every day of the week**, but this would be a jumbled mess with **7 lines**!
- We could use no visualization at all, and instead present the information in a table
- This is valid, but how can we make the table more interesting and usable?

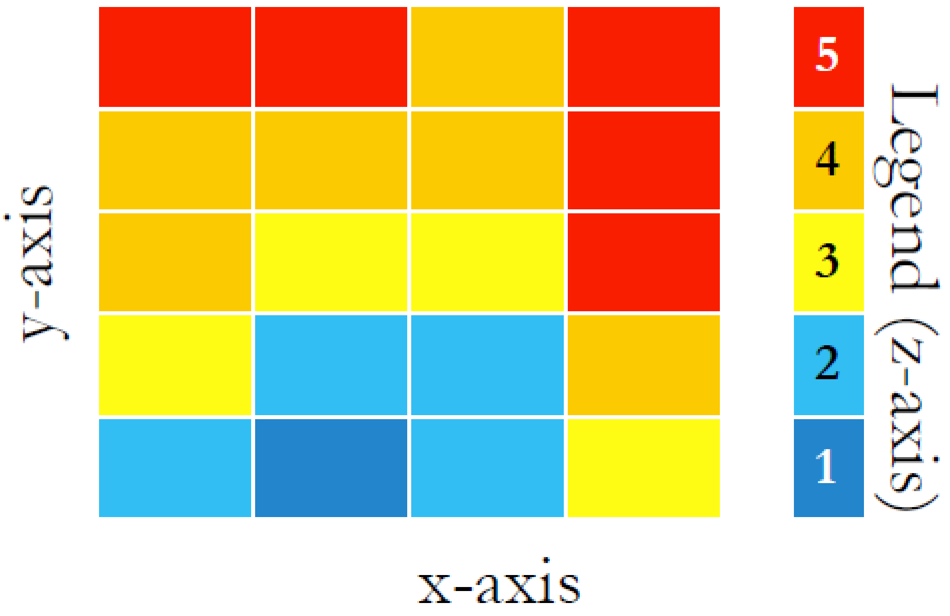
	MO	TU	WE	TH
03:00	34	32	31	…
04:00	15	24	22	…
05:00	22	10	33	…
06:00	13	14	19	…
…	…	…	…	…

Predicting Crime Hot Spots

Dynamic Crime Index: time-based pattern

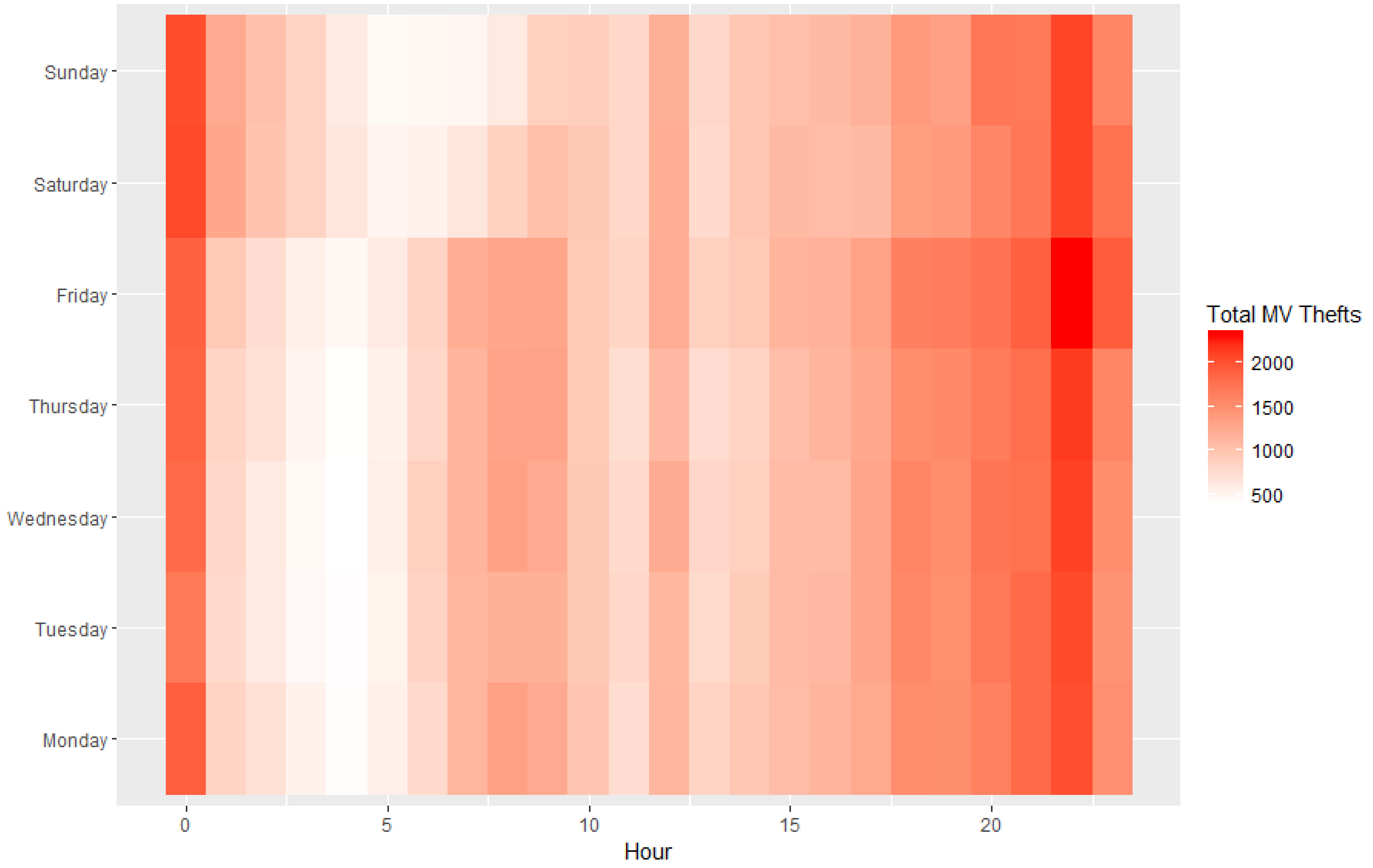
- **Heatmaps** are a way of visualizing data using three attributes. The **x-axis** and **y-axis** are typically displayed horizontally and vertically

- The **third attribute** is represented by shades of color. For example, a **low** number might be blue, and a **high** number might be red



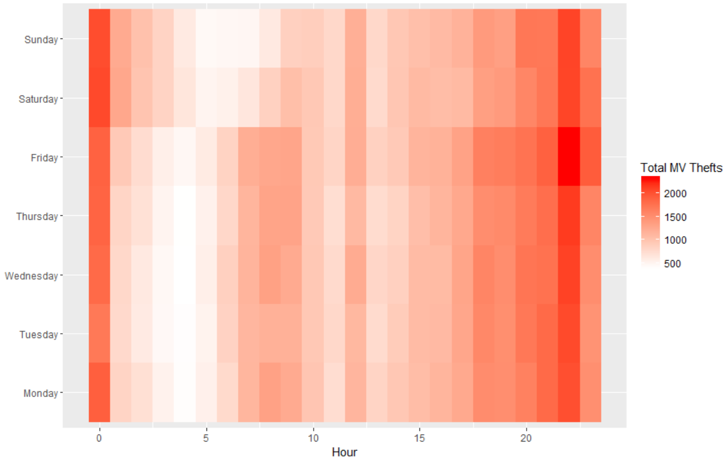
Predicting Crime Hot Spots

Dynamic Crime Index: time-based pattern

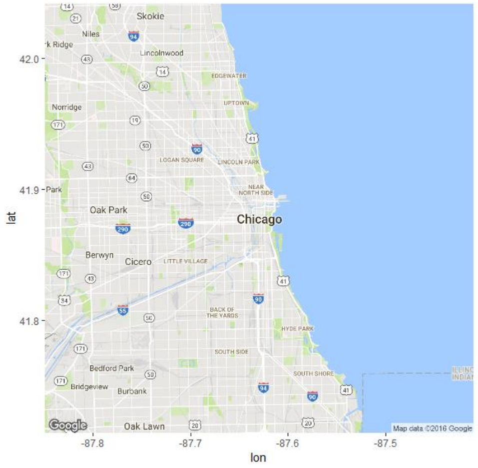


Predicting Crime Hot Spots

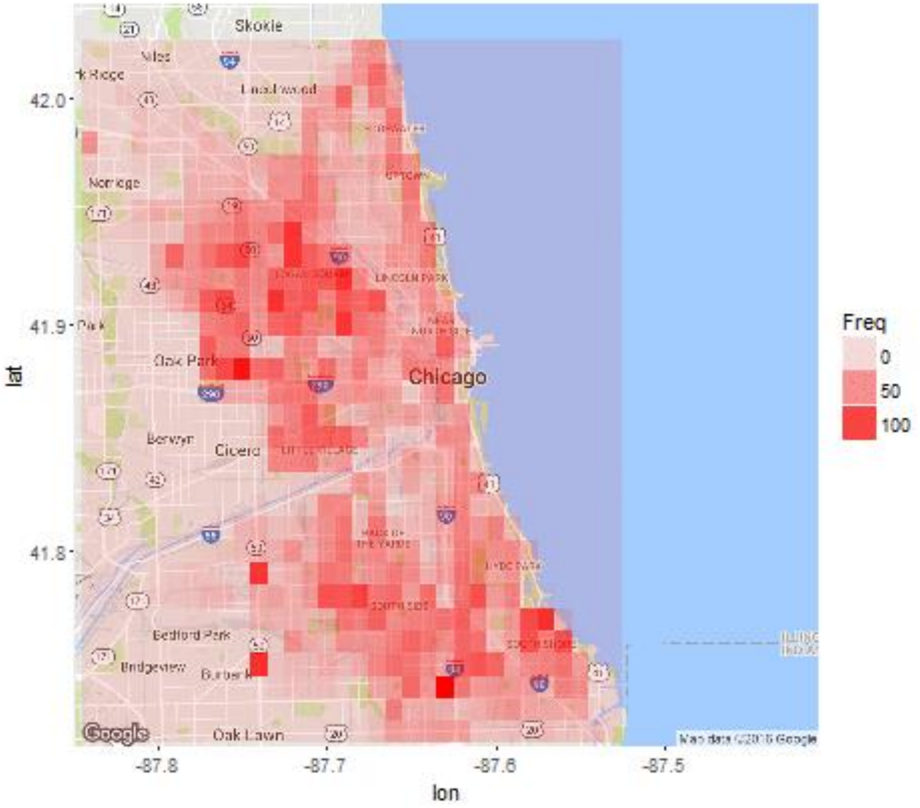
Dynamic Crime Index: spatial-temporal pattern



+

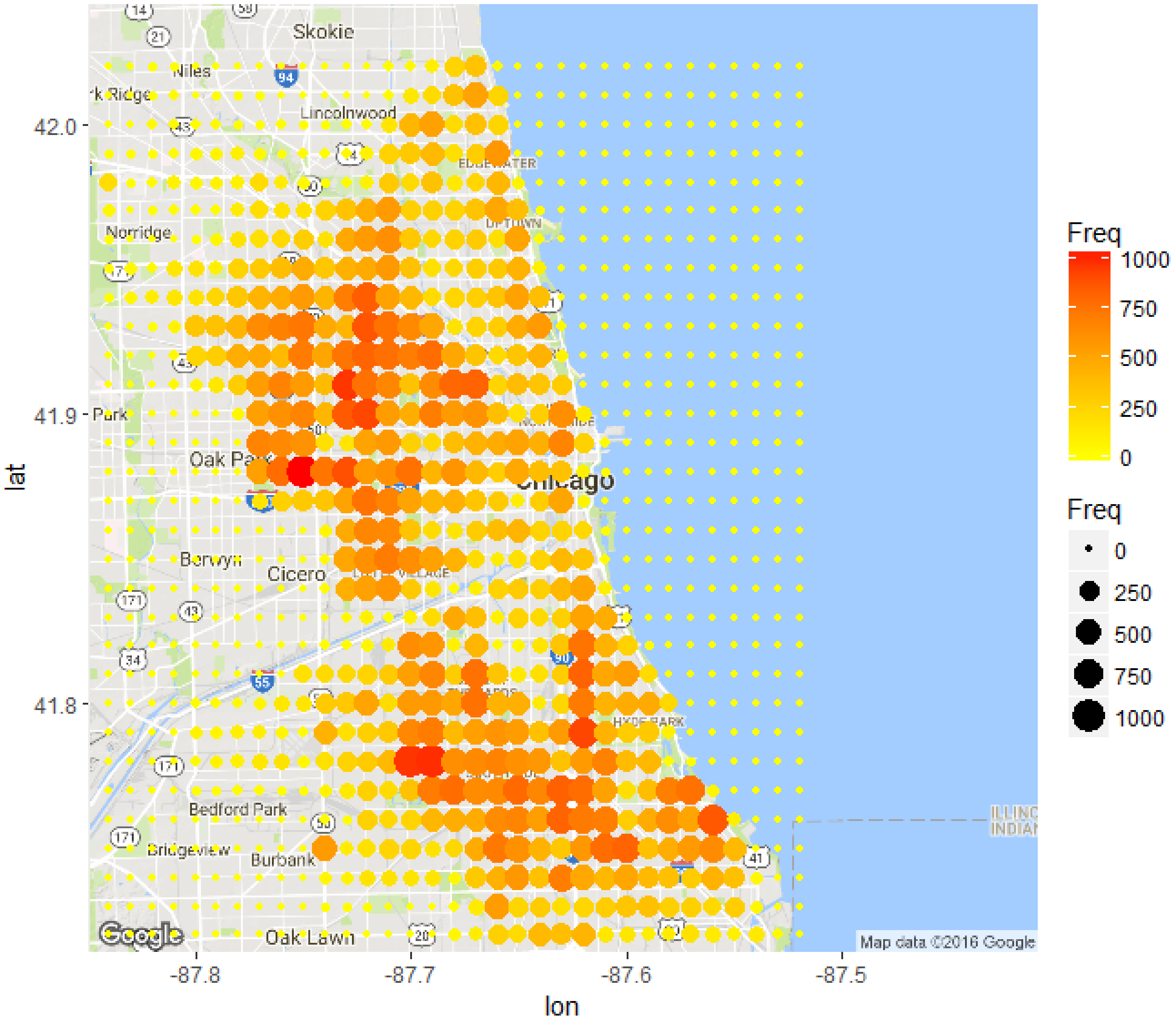


=



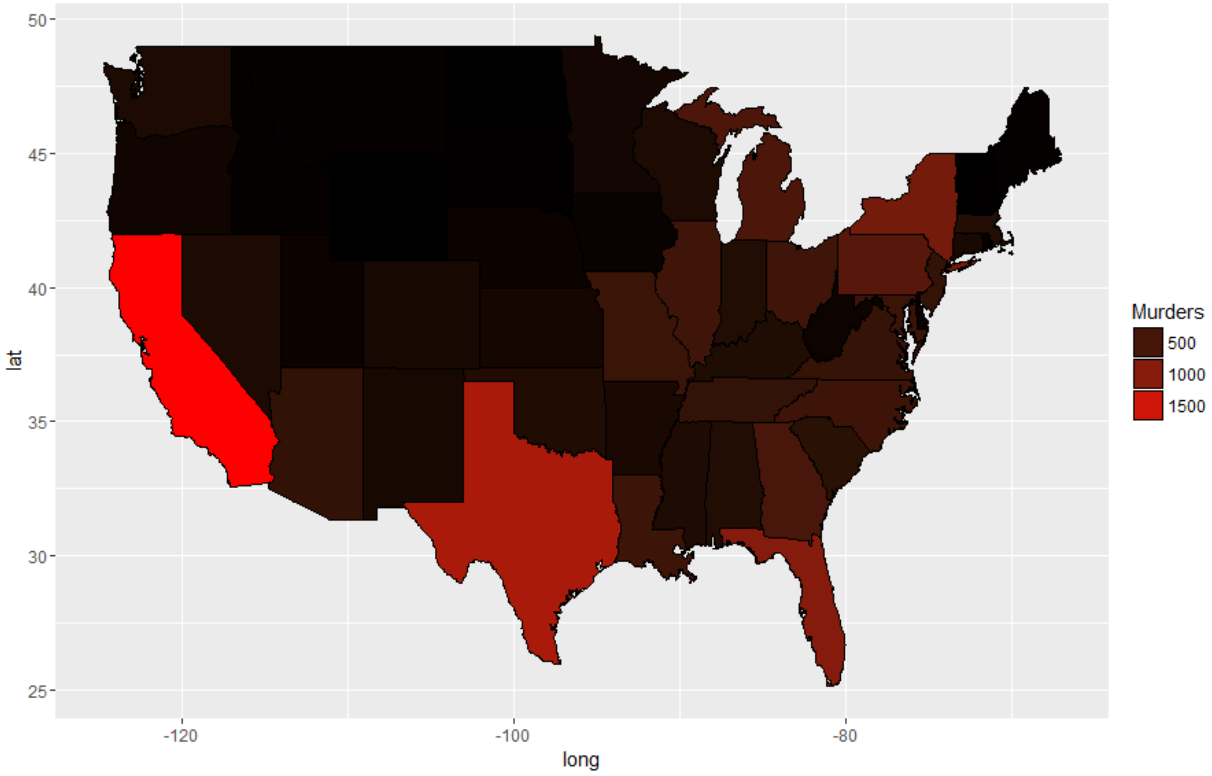
Predicting Crime Hot Spots

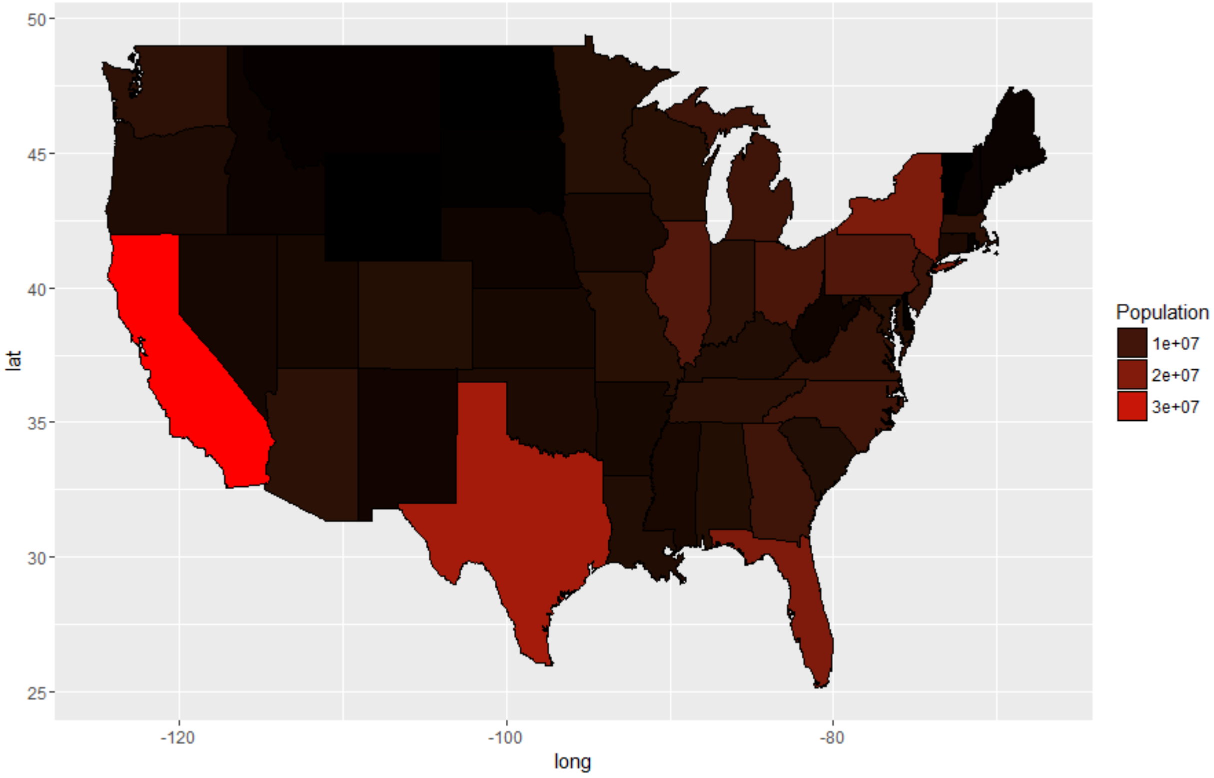
Dynamic Crime Index: spatial-temporal pattern



Predicting Crime Hot Spots

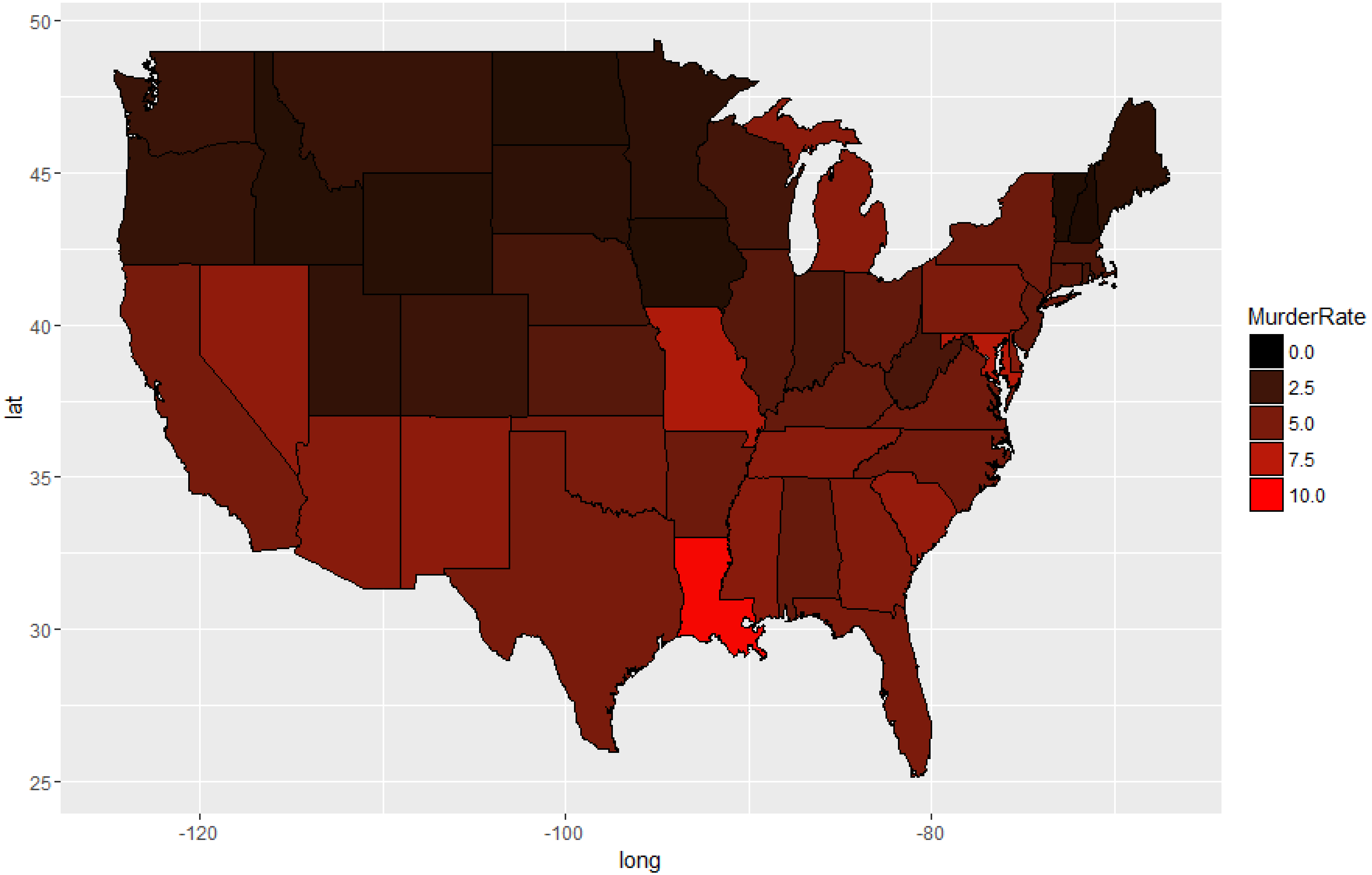
Murder pattern in the U.S.





Predicting Crime Hot Spots

Murder pattern in the U.S.

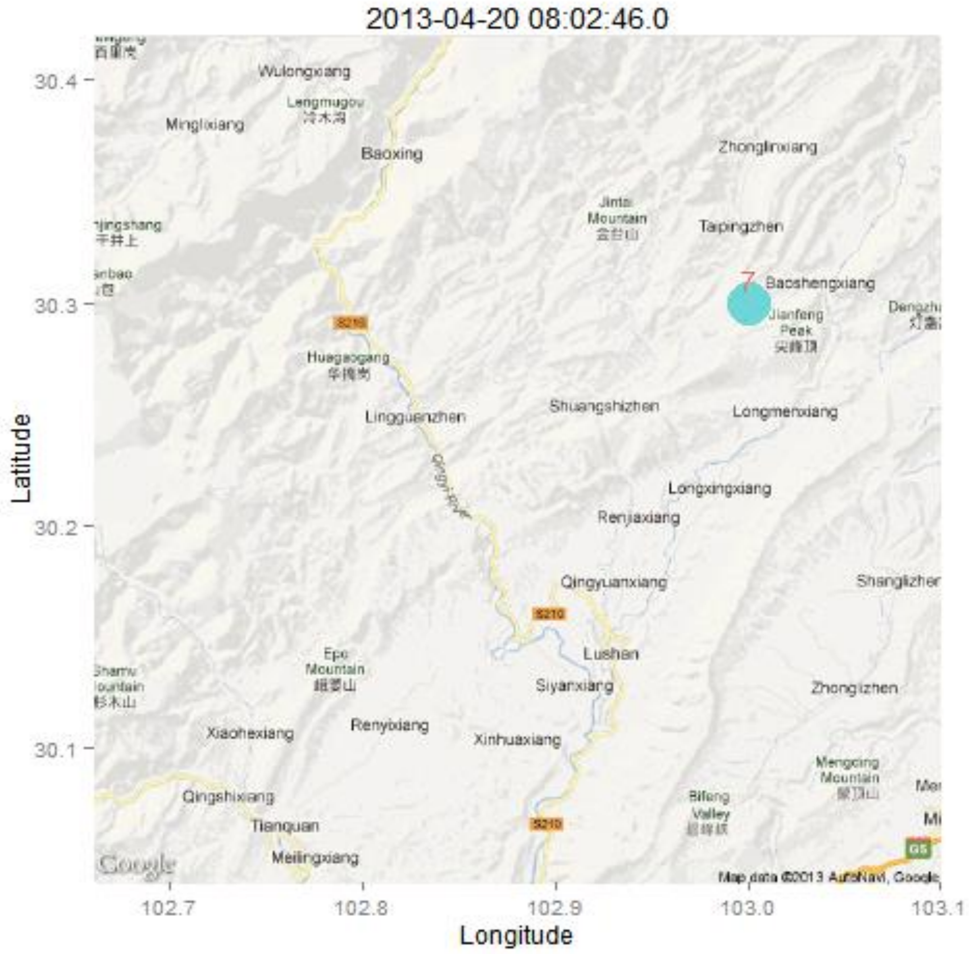


Predicting Crime Hot Spots

Dynamic Crime Index: a self-exciting point process model

Earthquake:

- Accompanied with a sequence of aftershocks;
- Hard to predict the initial earthquake;
- Aftershocks can be captured, which are influenced by time, location etc.



Predicting Crime Hot Spots

Dynamic Crime Index: a self-exciting point process model

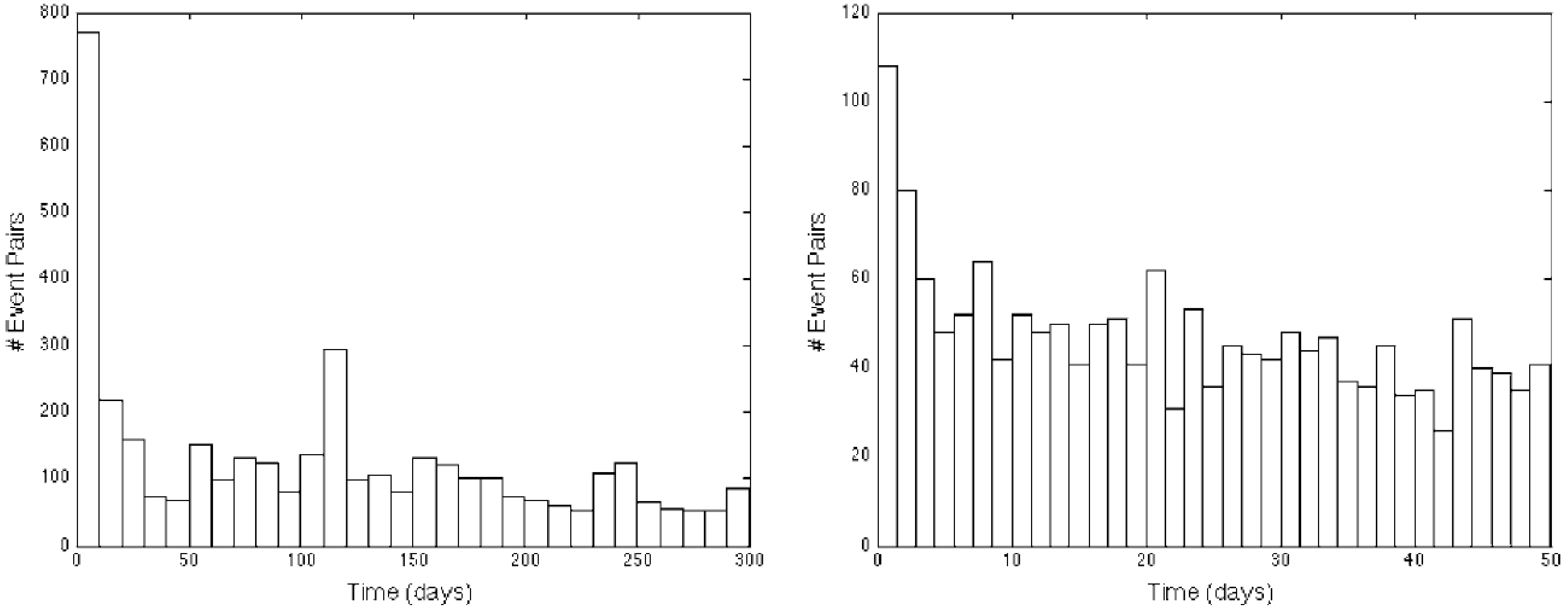


Figure 1. On the left, histogram of times (less than 300 days) between Southern California earthquake events of magnitude 3.0 or greater separated by 110 kilometers or less. On the right, histogram of times (less than 50 days) between burglary events separated by 200 meters or less.

Predicting Crime Hot Spots

Dynamic Crime Index: a self-exciting point process model

Given point data $(t_k, x_k, y_k)_{k=1}^{N}$ and a self-exciting point process model of the form,

$$\lambda(t,x,y) = \nu(t)\mu(x,y) + \sum_{\{k:\, t_k<t\}} g(t-t_k, x-x_k, y-y_k), \quad \text{(A.1)}$$

we iterate the following until convergence:

Step 1. Sample background events $\{(t_i^b, x_i^b, y_i^b)\}_{i=1}^{N_b}$ and offspring/parent interpoint distances $\{(t_i^o, x_i^o, y_i^o)\}_{i=1}^{N_o}$ from P_{n-1}.

Step 2. Estimate ν_n, μ_n, and g_n from the sampled data.

Step 3. Update P_n from ν_n, μ_n, and g_n using (8) and (9).

Assuming model correctness, the probability that event i is a background event, p_{ii}, is given by

$$p_{ii} = \frac{\mu(t_i, x_i, y_i)}{\lambda(t_i, x_i, y_i)} \quad (8)$$

and the probability that event j triggered event i, p_{ji}, is given by

$$p_{ji} = \frac{g(t_i - t_j, x_i - x_j, y_i - y_j)}{\lambda(t_i, x_i, y_i)} \quad (9)$$

Predicting Offenders or Victims

Node similarity:

- observed attributes of nodes;
- difficult to obtain the contents and attributes of nodes

Structural similarity

- Structure of networks (communication, geographical information etc.)
- the structural information is easier to get, and more reliable to trust

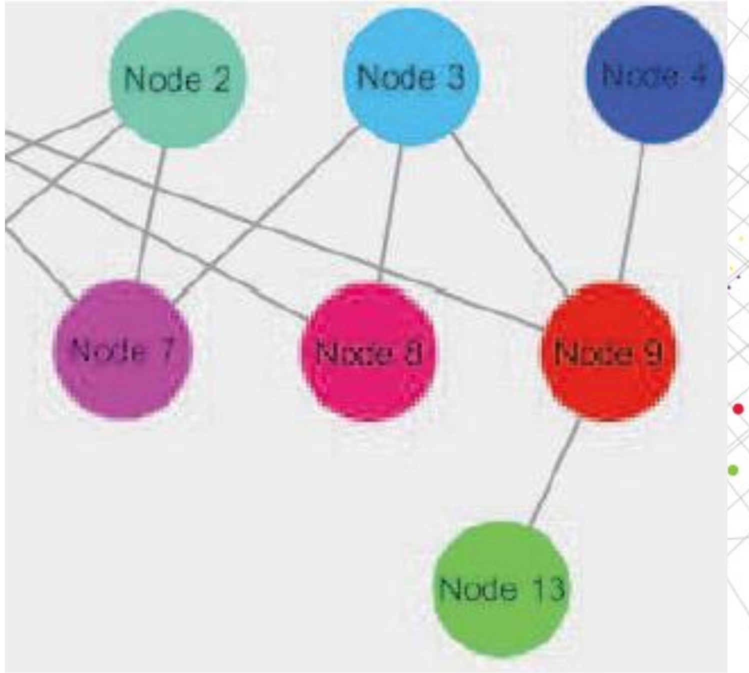
Predicting Offenders or Victims

Structural similarity index: distance between the nodes

Considering resource allocation (RA) index, although it could give good predictions for networks with large clustering coefficient, it performs badly for sparse networks. Notably, RA index is based on a simple assumption that each transmitter of a pair of indirectly connected nodes averagely distributes the resource receiving from one node to all the transmitter's neighbors.[18] According to this assumption, the similarity S_{ij} between two nodes v_i and v_j, which are not directly connected, can be defined as the amount of resource that v_j receives from v_i. Particularly,

$$S_{ij} = \sum_{z \in \tau(i) \cap \tau(j)} \left(\frac{1}{k_z} \right), \tag{1}$$

where $\tau(i)$ denotes the set of neighbors of v_i, $\tau(j)$ denotes the set of neighbors of v_j, k_z denotes the degree of node v_z, which is the common neighbor, namely, the transmitter of node v_i and v_j.



CRIME FORECASTERS

Police are turning to big data to stop crime before it happens. But is predictive policing biased—and does it even work?

By **Mara Hvistendahl**, *in Pittsburgh, Pennsylvania; Photography by* **Stephanie Strasberg**

"They're not predicting the future. What they're actually predicting is where the next recorded police observations are going to occur."

William Isaac, the Human Rights Data Analysis Group

"There are some cities where they have done a great job on hot spot policing, and they have terrible relationships with their communities of color."

Cameron McLay, Pittsburgh Bureau of Police