Detecting Crime Patterns

May 9th, 2019

Lori Pollock
Chief of Crime Control Strategies

Evan Levine
Assistant Commissioner of Data Analytics
What is a crime pattern?

“A series of crimes committed by the same offender or group of offenders…
To identify true patterns, one would need to consider information beyond simply time and space, but also other features of the crimes, such as the type of premise and means of entry.”

- Wang et al 2013
Organizational Challenge

Tens of thousands of complaints reported each year. When a new complaint is reported, analysts in each precinct are tasked with identifying related complaints. These patterns, after review, inform NYPD operations.
Patternizr helps us meet this challenge

• Designed for robberies, burglaries, and grand larcenies
• Built into the Domain Awareness System
• Patternizr helps the analyst make decisions, it does not make decisions for the analyst
Supervised Machine Learning

Historical patterns
(~10,000 of each crime type)

Training patterns

Validation patterns

Complaint pair similarities

Features extracted from complaint pairs

Random forest generation
What is a Random Forest?

Derived from classification trees which rely on features.
For example:

```
Round?  
  Yes    No
Apple   Yellow?
  Yes    No
Banana  Carrot
```

Patternizr uses features to classify complaint pairs. The output of Patternizr is the probability that a pair of complaints are in a pattern together.
P.O. records information about the crime:
• Time of occurrence
• Place of occurrence
• Premise type
• Property taken
• Suspect description
• Perpetrator’s statements
• Narrative of what happened
• Etc.
What features does Patternizr use?
Deployment and Production Process

1a. Train algorithm
1b. Test accuracy and fairness
1c. Prune features
2. Perform historical load
3. Integrate into Domain Awareness System
4. Train analysts
5. Perform differential loads and monitor usage

Iterate
Approximately 1/3 of test patterns perfectly rebuilt, and approximately 4/5 at least partially rebuilt.
Efficiency of Patternizr vs. a simple baseline
Safeguards ensure Patternizr is used fairly

- Sensitive attributes were hidden from the algorithm.
- Outputs of the algorithm were tested for fairness.
- Several layers of analytic and supervisory review required.