A Citizen's Guide to HunchLab

Information about what HunchLab does and related data

DRAFT Last Updated: 2017-07-11

HunchLab

Many residents and advocates are taking an increased interest in how policing practices impact our communities. At the same time, police departments are adopting new technologies that increase their effectiveness. One area of particular interest is in crime prediction or so-called “predictive policing.” Our HunchLab product, for instance, provides geographic crime forecasting that is used to allocate police patrols -- helping police commanders to know where and when crimes are most likely to occur.

Questions arise about how a system like HunchLab works, what types of data we hold on behalf of a police department, what data we generate while the system is being used, and what may be the positive and negative consequences of adopting such a system. We believe that transparency and fairness is critical in the application of algorithms within civic processes. Our hope is that this document helps you to understand more about HunchLab and its potential for making an impact in your community.

Sincerely,
Jeremy Heffner
Product Manager, HunchLab

Introduction

HunchLab is a patrol management system designed to make officers more effective in reducing overall reported harm in a community through proactive patrol. For many years, police departments have deployed patrols based upon hotspot maps of past crime locations assuming that future activity will likely occur at the same places. Instead of assuming that past crimes are the best predictor of future crimes, systems like HunchLab use machine learning algorithms so that the computer can learn how best to estimate the risk for future crimes at a particular location. These forecasts or predictions are then turned into a patrol plan and disseminated to officers in the field. Because such systems influence the day-to-day operations of a police department, it is important that such systems help to align police activity to address the problems that are affecting a community. In the next section we provide a quick snapshot of the high level items you should know about HunchLab. In subsequent sections, we outline more details about the main components of the system that influence the activity that HunchLab recommends to the police.

What You Need To Know (In Three Pages)

There are two main types of systems labeled as “predictive policing”. One type makes predictions about individuals likely to commit crimes. The other type makes predictions about places where crimes are likely. Currently, **HunchLab only makes predictions about places**.

HunchLab’s goal is to **design a patrol strategy to minimize the overall amount of preventable reported harm**:

- **Harm**: By harm we mean turning crime counts into meaning: how important is it to prevent a burglary versus a homicide? These are value decisions that HunchLab needs to know to balance focus between different crime types. We need to also think about the harm that can be created by police actions and choose patrol strategies and tactics that aim to minimize created harm.
- **Reported:** Police departments are only aware of crimes that are reported to them. Some argue that this bias makes crime forecasting impossible or unfair, but we believe that police activity should reflect what the community is reporting as problems. Such meaningful bias in the data helps align police activities with what the community cares about. There are of course a few exceptions to this principle; for example, when a victim doesn’t report crimes due to fear or shame. If, instead, reporting biases are due to distrust of the police, then we believe that letting the bias exist within the data is appropriate.
- **Preventable:** Some crimes are very targeted while others are opportunistic. The nature of some crimes makes it harder to prevent them (events that happen indoors, for example). Estimating how preventable a type of crime is helps to align resources with where they will do the most good.
- **Patrol Strategy**: A patrol strategy determines where officers spend time. Some officers respond to emergency calls and only have short periods of time. Other officers are dedicated to proactive work and may have an entire shift free.

To design a patrol strategy that meets these goals, HunchLab makes crime forecasts for individual types of crimes. For example, a specific police department may use HunchLab to model burglaries, aggravated assaults, robberies, homicides, motor vehicle thefts, and thefts from motor vehicles. **A separate machine learning model is created for each crime type**. Machine learning is where the computer uses example data points to learn how to solve a task. In this case, the task is to predict the level of reported crime for a small area across a few hours. The computer has access to a few types of data to solve this task. It knows the history of reported crime at that location and nearby. It also knows information about the location: How close to a

bar or bus stop is it? Is there a highway on-ramp nearby that lets offenders escape easily? It also has information about the time period: Is it a Friday night just after bars close? Is school in session today? The computer uses these data points to design rules that separate high risk conditions from low risk conditions. Although multiple data sets are used in HunchLab, It's important to think about what isn't being used. **HunchLab doesn't have information about specific individuals; it doesn't use the content you post online such as on Facebook or Twitter to make predictions.** The data we use is honestly quite simple and often publicly available.

If you are concerned about tools like HunchLab being used by your police department, here are some questions that you should ask. For each question we outline why you should ask it and what the answer for HunchLab would typically entail.

- **What are the overall goals for adopting the system?**
    - **Why**: The adoption of any system should have goals in mind; taxpayer money is being spent on such an effort so it should have a purpose.
    - **HunchLab**: Police departments often tell us that their goal is to increase the impact of their patrol activities to reduce crime. Sometimes this is to be more effective with the few officers a department has; other times it is to address crime levels that are considered "high".
- **What was the process before HunchLab? How will police actions change?**
    - **Why**: Sometimes people are concerned about biases in policing being perpetuated with new algorithm-based systems. If this is a concern of yours, you should think about how a new system will differ from what it replaces and whether it helps or hinders addressing such issues.
    - **HunchLab**: Before using HunchLab, most police departments design patrols based solely upon hotspot maps of past crimes. Such maps are often updated every few weeks and are created on the assumption that past locations of crimes will forecast future locations of crimes. Instead, HunchLab learns what matters in anticipating crime events. Patrol activities are then metered out in response to this risk.
- **What crimes are being modeled? What value decisions are being made? Who is making those decisions?**
    - **Why**: Police fulfill the role of serving the public to maintain safety. As such, the activities of the police should meet the definition of safety of the community they serve. There are value-based decisions in this process and knowing who is influencing that process is important.
    - **HunchLab**: Most clients use HunchLab to model major crimes like burglaries, robberies, assaults, homicides, and motor vehicle thefts. We recommend that police departments use external assessments of harm in setting the value

weightings in the system. That process could be a stakeholder group or published research into the impact of different crimes on the community.

- **What data is being used by the system?**
  - **Why**: Knowing what inputs are being used by a system helps you to understand how it works. It also helps you to advocate for changes to improve its fairness.
  - **HunchLab**: Different 'crime' data is appropriate in different situations and every client will have a slightly different list of data. Typically crime reports drawn from a records management system or calls for service drawn from a computer aided dispatch system are used. Other data utilized is often publicly available and so disclosing its use in such a system is straightforward.
- **What potential harms could happen? How are you being proactive to address them?**
  - **Why**: It's important to think about the risks for any new endeavor. Putting into place mechanisms to proactively assess the effect of changes will help to catch such issues early.
  - **HunchLab**: We want to prevent an algorithm such as HunchLab from increasing bias so we prevent the system from using data representing activities generated by the use of the system to prevent feedback loops. Additionally, we probabilistically select locations for patrols to reduce over saturation at specific locations.
- **What do officers see / how do they use it?**
  - **Why**: Seeing something first hand often helps to understand it better. Asking to see what it looks like is simple to do.
  - **HunchLab**: Webinars that we produce for our clients are posted on YouTube which will give you a sense of what the officer sees in the field. Additional information is available on our resources page: https://www.hunchlab.com/resources

How HunchLab Works In More Detail

Forecasting Crime

At the core of HunchLab is a crime forecasting engine. These forecasts or predictions are used to allocate patrol resources. The forecasting engine uses ensemble machine learning approaches that can incorporate the following crime patterns into a single prediction of criminal risk:

- Baseline crime levels
    - Similar to traditional hotspot maps
- Near repeat patterns
    - Event recency (contagion)
- Risk Terrain Modeling
    - Proximity and density of geographic features (points, lines, and polygons)
- Routine activity theory
    - Offender: Proximity and concentration of known offenders
    - Guardianship: Police presence (historic AVL / GPS data)
    - Targets: Measures of exposure such as population, parcels, or automobiles
- Collective Efficacy
    - Socioeconomic indicators, heterogeneity, etc.
- Temporal cycles
    - Seasonality, time of month, day of week, time of day, etc.
- Recurring temporal events
    - Holidays, sporting events, etc.
- Weather
    - Temperature, precipitation, etc.

Our belief is that the use of non-crime data sets as variables within a crime prediction system is important, because variables based solely upon crime data become skewed as predictions are used operationally. For instance, as crimes are prevented in mission areas due to police response, the only variables identifying areas as high risk are skewed in other systems. By including other data sets, our system is more robust against this issue.

In some ways, the model building process in HunchLab mimics the thought process of an experienced crime analyst. For instance, consider asking an analyst to decide where to place patrol resources for a given upcoming time period. She may start by looking at where crimes have occurred in concentration previously and delineate hotspots of activity. Based on her past

experience, she may know that during this particular time period, schools dismiss their students, which increases petty crimes around the schools in the neighborhood. She builds up many such layers of knowledge and balances these various concerns to form a plan. After the time period concludes, she may go back and look at where activities occurred to see if she can determine additional insights into the crime patterns to include in future plans. HunchLab incorporates machine learning concepts to help the software "think" like a crime analyst by imitating years of experience drawn from a police department's own data.

The concept of machine learning is to teach a computer to accomplish a particular task. In this case, we want to teach the computer to determine how likely a particular crime type is to occur at various locations for a given time period. We start this process in HunchLab by forming a set of training examples using the past several years of crime data. Each training example contains the theoretically derived variables we explained above, as well as the outcome (how many crime events occurred). For an entire municipality this training set will often include many millions of example observations. We can then start building the model.

The primary model HunchLab currently uses is a stochastic gradient boosting machine (GBM) comprised of decision trees trained using the AdaBoost loss function implemented using the open source GBM package. This model is built to forecast whether one or more crime events will occur in a given space-time raster cell (a binary outcome). The general way this model works is as follows:

- Begin by selecting a random portion of the training examples.
- Build a decision tree that separates examples of where crimes occurred from ones that didn't based upon the variables.
    - For instance, the first decision within the tree might be interpreted as: "if no event happened in the last year in this location, it is very unlikely for a crime to occur today". The decision tree then splits the examples into two sets: (1) where a crime occurred during the past year and (2) where no crimes occurred during the past year.
    - Within each set, the process repeats. For example, the next decision for the set of locations with crimes in the past year might be interpreted as: "if an event happened in the last week, it is more likely for one to occur today". This set of examples would again be split based upon this decision rule.
    - This process continues to build out a decision tree that describes why crimes occur where they do.
- The decision tree is then used to make predictions of how likely crimes are for each observation in the entire training data set.
- This completes one training iteration within the boosting machine.

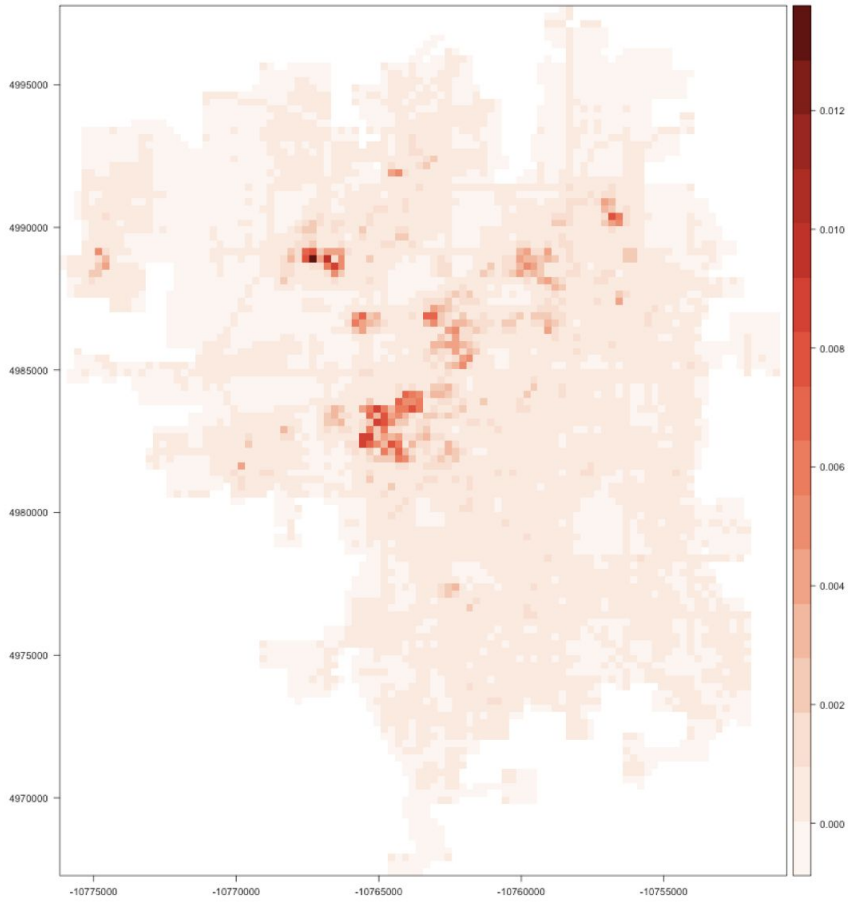
The modeling process then begins again:

- We start by selecting another random portion of the training examples. This random sampling process is why the model is stochastic.
- In this next iteration, we build another decision tree (in the same manner as above). This time, however, we build the tree to predict the errors from applying the first decision tree model to this new sample of observations. In this way we are attempting to correct our mistakes. This concept is called boosting.
- We then use these two trees to make predictions across the entire data set.
- As we conduct this process, we can keep track of how many training iterations within the machine have made incorrect predictions for each training example. We increase the importance (via weights) of observations that we continue to get wrong and decrease the importance of observations that we continue to get correct. This process is called adaptive boosting (AdaBoost).
- When we build the next decision tree, we tell it to focus on the observations that we continue to get wrong via these weights.
- Training iterations continue several hundred times. The resulting model represents tens of thousands of decision rules of why crimes occur where they do.

We conduct this entire process several times, each time holding back a portion of the example data. We can then use each of these models to make predictions for this held-out set of data to see how accurate the model is when we apply different quantities of training iterations from the model. For instance, if the models have 100 training iterations, we may find that the most accurate predictions come from only using the first 53 iterations. This process is called cross-validation and prevents our models from overfitting the training data. Finally, we begin the entire process again using the whole data set to build a model with the correct number of training iterations. In this example, we would use 53 iterations.

As you can see, this modeling process mimics some activities that an analyst would go through in making decisions of where to focus resources. The predictions from this model are whether one or more crimes will occur or not. We then need to translate these probabilities into expected counts. We do this by calibrating our predictions using a generalized additive model that assumes a Poisson distribution and is trained using the open source MGCV package. This regression model both translates the outputs of our model to expectations and calibrates the predictions. For example, the above model might slightly over-predict crimes on Tuesdays. This calibration step would lower the predictions for Tuesdays to center them on the training data. The process of using one model's outputs as another model's inputs is called model stacking. These models are then saved and used to generate predictions.

The predictions are calibrated count expectations for each raster cell for a given period of time. You may picture predicted counts to be numbers such as 0, 1, 2, or 3. In actuality, the predictions are real numbers that are often fractions such as 0.000001, 0.02142, or 0.12482. This represents the fact that the nature of crime is such that no software solution can say that a crime is going to happen at this specific corner at this precise time. For a small raster cell and time period, it is almost always more likely that no crime will occur. What is important is that we can use these predictions to measure the relative risk of events between locations, time periods, and different crime types, so that we focus on the most likely types of events at the most likely locations and times.



Aoristic Times

Often the exact time that a crime event occurred is unknown. A prime example of this phenomenon is residential burglary, where the homeowner discovers the burglary when they arrive home from work. We handle this phenomenon by using aoristic time ranges within our model. For each crime event imported into HunchLab, a time span during which the event

occurred (sometime between 9:00 AM and 5:00 PM, perhaps) is provided. If the time of the event is certain, these start and end times are simply set to the same time. When HunchLab builds the training data set that is used in the modeling process described above, we use these time ranges to determine the possible outcome scenarios for the observation. For instance, assume we are creating a training example for a particular raster cell from 9:00 AM to 10:00 AM on January 16, 2014. There is one crime that may have occurred during this period. The event happened between 9:00 AM and 11:00 AM. We therefore have two scenarios from 9:00 AM to 10:00 AM: (1) 0 events occurred and (2) 1 event occurred. We assume that the event is uniformly likely to have occurred during the aoristic time frame (2 hours in length), so each of these scenarios is equally likely (1 hour / 2 hours = 50% probability). Both scenarios are placed into the training example set for use in modeling with a weight of 0.50 for each scenario.

As the modeling process selects random portions of examples for training during each training iteration, it will sometimes include the first scenario and sometimes the second one. This approach nicely represents the relative uncertainty of the exact event time.

Model Variations

There are several parameters that define the exact functionality of the model building process described above. HunchLab can also use varying amounts of historic data in building the model in order to balance the desire to have more examples with the desire to use recent examples. We adjust some of these various parameters to generate a few variations of our modeling process. The system automatically scores each variation on a held-out data set (the most recent 90 days of crimes, for instance) and then selects the best performing variation for use in production.

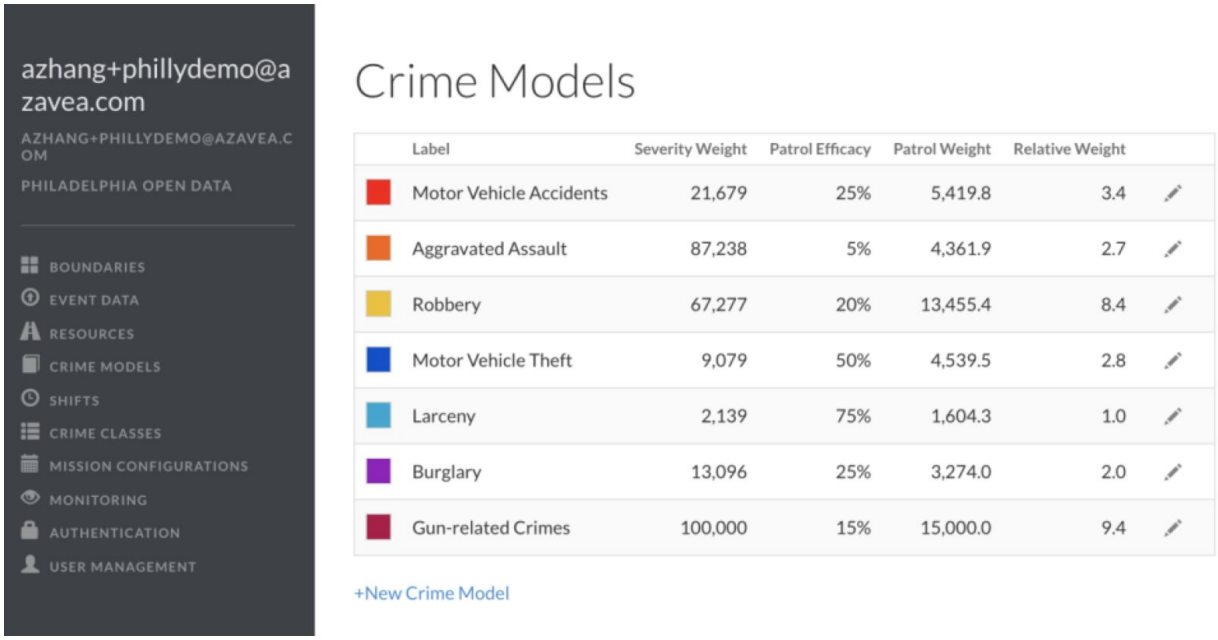
Calculating Preventable Harm

Once we have forecasts for individual types of crimes we then need to turn these numbers into a patrol allocation. One approach is to treat all types of crimes as equivalent and try to reduce the overall count. To our knowledge this is how all other forecasting software works, but it does not help to align policing activity with making the most impact. High volume crime types will heavily influence the recommendations even if they are less severe events. To address this we weight the forecasts with two values:

- **Severity Weight**: Severity weights enable HunchLab to know how important it is to prevent each type of crime. As the severity of a crime model increases, the system will place more emphasis on that crime model relative to the other models. HunchLab recommends setting severity weights to align with the societal impact of crime. Examples of severity weight values to use include the RAND Corporation's Cost of Crime.

- **Patrol Efficacy**: Patrol efficacy is another setting which tells HunchLab how much emphasis to place on a type of crime. HunchLab is a directed patrol tool. Because of this, HunchLab should direct officers to focus on types of crime that are more preventable by directed patrol. The higher the patrol efficacy is set for a given type of crime, the more prevalent that crime will become in mission generation. There is less literature on the specific numeric efficacy of patrol for certain types of crime, so this value can be an opportunity for departments to use local patrol knowledge to influence the importance HunchLab assigns a given type of crime.

Here is an example of how the configuration page looks for these values looks within HunchLab:



Label	Severity Weight	Patrol Efficacy	Patrol Weight	Relative Weight
Motor Vehicle Accidents	21,679	25%	5,419.8	3.4
Aggravated Assault	87,238	5%	4,361.9	2.7
Robbery	67,277	20%	13,455.4	8.4
Motor Vehicle Theft	9,079	50%	4,539.5	2.8
Larceny	2,139	75%	1,604.3	1.0
Burglary	13,096	25%	3,274.0	2.0
Gun-related Crimes	100,000	15%	15,000.0	9.4

The result of this weighting process is an assessment of the predicted preventable reported harm in each area of the jurisdiction. The next step is then to decide how to select locations using this information.

Fair Allocation

A straightforward way to select locations for patrols is to rank locations based upon risk from high to low and then select the highest risk locations. This is how HunchLab originally worked, but there are several downsides to this approach. Risk locations tend to persist over time, so the same locations were frequently selected. Officer's day to day work can then become monotonous and suffer from the sense that nothing is changing. At the same time, the unintended side effects of saturating an area with police presence increases with the degree of

saturation. To solve these issues, we now recommend a probabilistic selection of locations within HunchLab using lottery weights. The process works as follows:

- Filter to the applicable choices based upon the geographic area desired (e.g. beat)
- Take the weighted risk values and z-score them. Values then represent the number of standard deviations above or below the mean.
- Filter out choices below a configurable threshold. For example, only keep choices where the risk is average or higher. This prevents us from selecting low risk locations.
- Raise the remaining scores to a configurable power. We typically use a value of 2 or 3 for the power. This process makes larger scores grow faster than smaller ones. If we raised the scores to a very very high power then the highest risk locations would almost always get selected which behaves like the deterministic selection we originally utilized. If we raised the scores to a power near zero then the system would select locations randomly.
- Use the values as weights in a lottery system to select locations.

Tactics

Why Incorporate Tactics in HunchLab?

- **Changing officer behavior** - A typical hot spot map gives officers large focus areas that do not change. With such generic guidance, officers tend to stick to the major roads and intersections within a hot spot in vehicle patrol, even though this is not always the most effective patrol strategy. HunchLab prompts change by showing officers changing tactics in small, clearly delimited areas that change every shift.

- **Community-specific** – The tactics that work in a large city may not work in a rural area because the drivers for crime are different. HunchLab can display tactics that command staff has learned over the years work in their own community, or that they have always wanted to have officers try. Tactics can be customized and updated at any time. The system then experiments with the entered tactics and measures which ones work best.

- **Evidence-based** – To use the right tactics, a department first has to determine what works. Randomized experiments are complicated and can take years to implement. HunchLab simplifies and automates the process of evaluating tactics. As the system learns which tactics are working it recommends those best tactics to officers more often, automatically incorporating evidence-based tactics into patrol routines and enabling flexibility.

- **Harm-conscious** – HunchLab helps departments to think about the total costs and benefits of tactics. Pedestrian stops may result in deterring crime, but at what cost to citizens? When determining what steps officers should take, departments must take into

account not only the potential to reduce crime but the potential harm of each method. By evaluating new tactics that reduce the potentially created harm in a community, police departments can advance their mission to be effective and fair.

Our Guidance On Crafting Tactics

HunchLab recommends that departments choose tactics rooted both in local knowledge and studied best practices. While every community is unique and require local understanding, there is a growing body of research around evidence-based policing that is broadly applicable regardless of place. Below we have outlined some initial questions departments may want to consider in choosing tactics.

Has the tactic been proven to work? There are a number of rigorous studies that have been completed in cities across the U.S. that explore questions of which police tactics actually work, which tactics produce negative consequences, and which tactics are promising but yet unproven. While it is certainly acceptable to try something new, existing research may shed light on the usefulness of related interventions which can inform the design of new tactics.

Some examples of evidence-based tactics include:

1) Hot spot foot patrol - In the Philadelphia Foot Patrol Experiment, officers patrolled in pairs by foot in violent crime hotspots for a three-month experiment phase. The randomized control trial found that in foot patrol treatment beats, one violent crime was reduced for every additional 4 arrests, 89 pedestrian stops and 8 traffic stops

2) Hot spot car patrol - The Minneapolis Hot Spot Policing Experiment found a 6 - 13 percent reduction in citizen crime calls and a 50 percent reduction in observed disorder in the hot spots that received additional patrol.

3) Problem-oriented policing strategies - Problem-oriented policing calls for police to identify the problems causing crimes and then use tactics designed to combat the specific problem. In a Glendale, Arizona study (White & Katz, 2013) police identified the problems leading to increased crimes at a local convenience store chain, including inadequate staffing and failure to respond to panhandling and loitering. Based on these issues, officers worked with store leadership and assisted in surveillance and enforcement. This led to a 42 percent crime decrease in targeted stores. In another example (Baker and Wolfer, 2003), police found that crime in parks stemmed from alcohol and substance abuse. Officers initiated a neighborhood watch group, patrolled to control alcohol and drug use, and enforced a new curfew law, resulting in reduced harassment and vandalism.

Officers can use a combination of patrol, victim-targeted interventions, and offender-targeted interventions. In the prior example, initiating a neighborhood watch group with monthly police meetings is an example of a victim-oriented tactic as it helps educate victims so they are less likely to be suitable targets. On the other hand, targeted enforcement of alcohol or drug use in parks is an example of an offender-oriented tactic. With HunchLab missions, it is important to think about designing tactics that can be done in short doses (10 to 15 minutes) in a small geographic region (several hundred feet by several hundred feet),

What may be the unintended consequences? Some tactics are not just ineffective, but produce negative consequences. Others may be effective, but have the potential to cause harm to the community that outweighs the positive benefit of crime reduction. Some questions to consider to avoid causing unintended harm include:

- Does this tactic hurt those it is trying to help?
- Does it punish innocent bystanders?
- Does it disproportionately impact minorities, or a specific demographic segment?
- Does it improve or worsen community relations?

Some tactics do not produce intended result, and instead cause harm. For example, Hovell et al. (2006) found that a Family Violence Response Team designed to reduce domestic violence was correlated with an increase in the likelihood of a treated family experiencing future violence.

Additionally, some effective tactics may have negative consequences that outweigh their positive benefits. In a study of eight cities, Wesley Skogan (2006) found that public encounters with the police carry a negativity bias; a negative police encounter has a 4 to 14 times higher impact on a citizen's perception of the police than a positive encounter. While the Philadelphia Foot Patrol Experiment showed that 80 stops led to one fewer violent crime, the number of potential negative experiences, and thus high cost of such a strategy, could outweigh the benefit of the deterred crime if the strategy is misused.

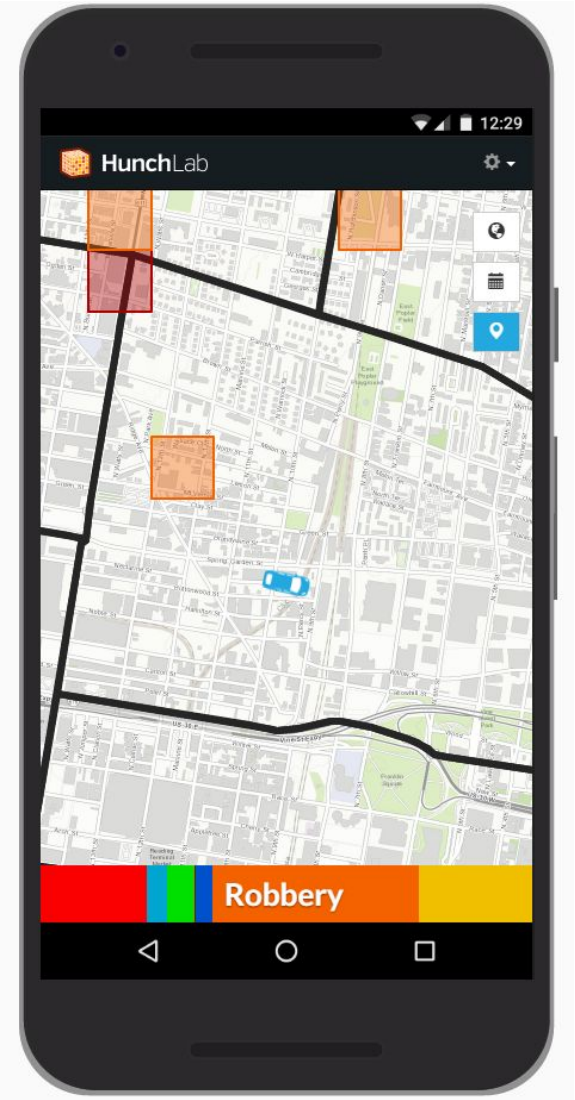
To minimize unintended consequences, HunchLab recommends doing thorough research of tactics before adopting them, following principles of problem-oriented policing, and thinking critically about how a tactic may impact certain demographic groups. In the case of untested tactics, HunchLab recommends focusing on proactive, specific, place-based strategies that try to minimize negative impacts. Existing policing tactics may inspire you to try variations that minimize such negative impacts. For example, many police departments stop vehicles for traffic violations and then determine whether the occupant has a warrant, etc. A variation of vehicle stops could be "vehicle stops with only warnings" meaning that minor traffic infractions (while the reason for the stop) would never result in a traffic ticket or citation but only a verbal warning. Such a tactic may help to limit the potential for disparate policing impact in minority communities, which have increased police presence due to crime unrelated to traffic issues.

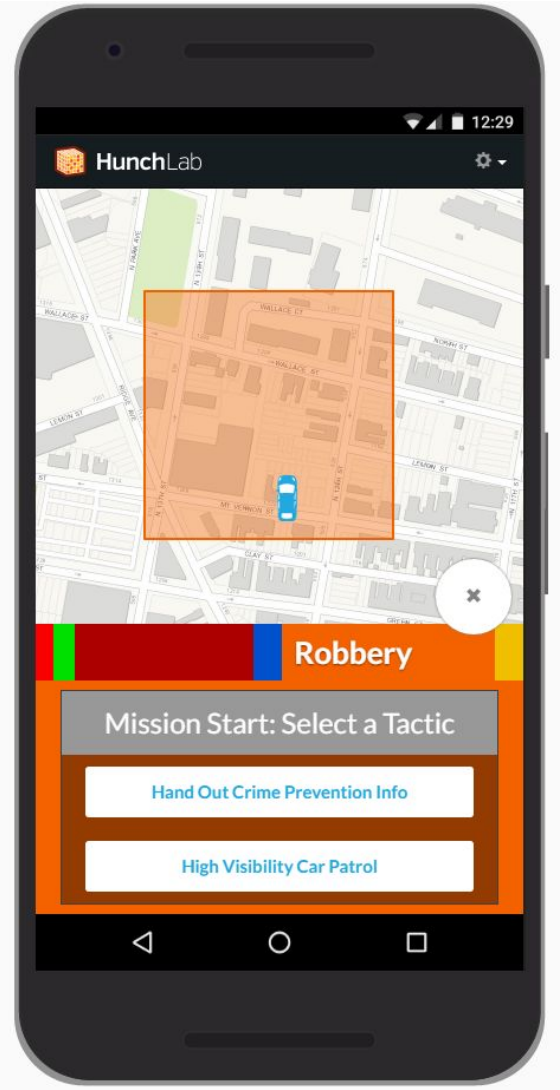
To add a more holistic, community-focused approach to policing, we also recommend using diverse tactics that involve community engagement. For example, if thefts from motor vehicle are a problem in your department, a department might enter three tactics: simple vehicle patrol, placing flyers on cars warning of the thefts, or speaking to employees at local businesses. Although all three tactics are a police response, two of the three tactics do not involve patrolling in the traditional sense and build positive interaction with the community. If a particular tactic is much more effective than the others it will be recommended more often by system. If, however, the tactics that are entered have similar results, the system will use them at the same rate. This presents a benefit in that applying a number of different tactics prevents a single tactic from always being utilized. That tactic may have unintended consequences that are hard to measure, so the mere fact that it is used less often due to more tactics being entered is a positive outcome.

How will the tactic be carried out? Once a department decides on a tactic, it can be carried out in a few ways. For example, a hot spot car patrol may be carried out with marked police cars or undercover cars. Again, turning to existing studies is a good place to start to understand the benefits and costs of particular options in unique cases. For general patrol strategies, HunchLab recommends instructing officers to patrol according to the Koper Curve Theory, which reduces the chance that particular areas will be over-policed and increases the effectiveness of patrol overall. Koper found that patrolling hot spots for 10-16 minute intervals, randomly every two hours resulted in the highest crime reduction benefit. After 16 minutes, the effective benefit of patrols decreased. Based on this theory, we have built a timer into HunchLab which shows officers a 15-minute timer to help them track how long they have been in a mission area and encourage them to move along to the next area.

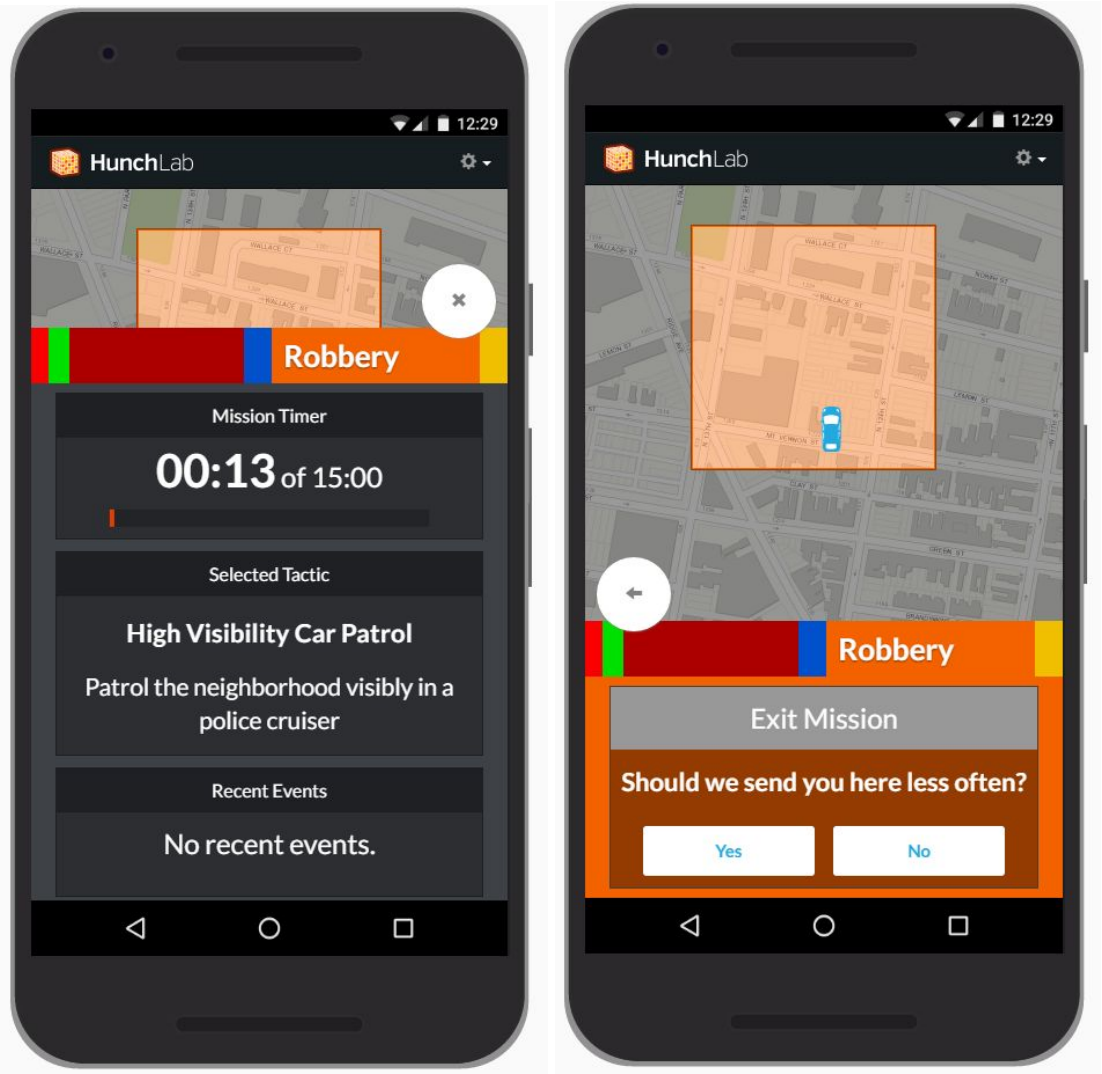
What Officers See

Officers access HunchLab through a web browser on laptops, tablets, and smartphones. The below scenario represents the typical workflow for an officer in the field using a smartphone. Picture that you are an officer assigned to a beat represented by the black boundary on the map below (left). The car icon shows your current location based upon the GPS chip in the device. You've just finished responding to an emergency call and now you head to the mission area represented by the orange box showing a focus on robberies. When you arrive, the application triggers you to select a tactic and you pick "High Visibility Car Patrol" (right).





The application then begins a timer to help you keep track of how much time you are spending in this location (left). You drive around the streets in the mission area keeping yourself highly visible. After 10 minutes have passed you click to complete the mission session. The system prompts you to answer a survey question to provide your feedback to improve the recommendations (right)



Throughout your shift you would repeat this process visiting mission locations in between other activities.

Data

To provide the HunchLab service there is a set of data that is loaded into the system from external sources. Additionally, there are data sets generated by HunchLab through its operation. The data within the system belongs to the client so we cannot release data without their permission. HunchLab asserts no proprietary rights over the data we manage on behalf of our clients. The following sections outline details about these data sets.

External Data

There are five types of external data used by the system: (1) data that represents events being forecasts such as crime incidents or emergency calls, (2) geographic data sets used as variables within the predictive model, (3) temporal data sets used as variables within the predictive model, (4) geographic boundaries delimiting parts of the jurisdiction, and (5) external GPS feeds used to measure patrol presence. Not all of these data sets will be utilized for a specific police department.

Event Data

This data represents the events being forecasted. These records are typically drawn from a police department's Computer Aided Dispatch (CAD) or Records Management System (RMS). The data is replicated to HunchLab. Each record has several attributes including: the geographic coordinates of the event, dates and times of when the event occurred, and the classification of the event.

Geographic Variables

This data is optional but most clients use geographic covariates within HunchLab. These geographic layers will represent features of the landscape such as the locations of bar, bus stops, highway onramps, population densities, etc. The data can be vector data such as points (bars), lines (railroad tracks), polygons (census data), or rasters (elevation data). HunchLab accepts ShapeFile format for these data sets.

Temporal Variables

This data is optional but most clients use temporal variables in HunchLab. The data is ingested in CSV format or via the HunchLab API. For a specific time period a specific value for the variable is presented to the system. For example, the temperature outside was 30 degrees between 1am and 2am today.

Geographic Boundaries

To allocate police patrols, the system needs to know how the police department organizes work. A typical setup will involve uploading multiple sets of boundaries into HunchLab. One boundary will represent the jurisdiction boundary. Other boundaries will break up the jurisdiction into the divisions, districts, sectors, or beats of that police department.

GPS Feeds

To measure police present, HunchLab sometimes ingests GPS feeds of police officers (cars or radios typically). GPS feeds can also be generated by the use of the HunchLab system on smartphones or tablets.

Configuration

Within the administrative interface to HunchLab, various configuration options are set. These options impact what event data is modeled, how important the different event types are, how to allocate patrol resources, etc.

In particular the following configuration data is present in HunchLab:

- Shifts
    - These entries divide the 24 hour day into distinct time periods. These periods will often align with police shifts but sometimes comprise parts of shifts. The shifts can be from 1 to 24 hours in duration.
- Resources
    - These entries define particular types of patrol resources and their patrol speed. For example, a car can patrol faster than a bike. The patrol speeds are utilized to determine how many areas should be recommended for that type of resource.
- Event Classes
    - These entries represent the distinct crime classes associated with imported event data. For example the external system may use a numbering scheme such as class '123'. In HunchLab that can be labeled as 'Robbery w/ Gun'. Crime classes are assembled into crime models.
- Event Models
    - Event models define the collection of events to forecast. For example an event model for assaults could be defined to only include the assault event classes that are relevant to patrols (excluding domestic assaults, for instance).

    - These records also define the 'severity weight' of the event. This value specifies how much harm is created when one of these events occur.
    - Additionally, the 'patrol efficacy' percentage is specified which represents how effective patrols are in addressing this event type.
    - Finally, these records define the text label and color used to represent the model.
- Tactics
    - These records represent the configured tactics that will display to officers within mission areas. Each tactic has a title and short description. The applicable crime models are selected so that specialized tactics that apply only to different types of crimes can be entered. Tactics can also be made active or inactive.
- Mission Configuration
    - A mission configuration defines how to translate forecasts into a patrol allocation. For example, if officers are assigned to geographic beats, then missions should be created for each beat area so that all officers have patrol recommendations.
    - The quantity of missions desired is also set either as a fixed amount or based upon available resource quantities. For example, if one patrol car is assigned to each beat, cars patrol at 5 miles per hour, the shift is 8 hours long, about 25% of the officer's time is free, and officers should spend 15 minutes in a mission 3 times during the shift, the correct quantity of missions can be created based upon these factors.
    - The allocation strategy is also defined in the mission configuration. One option is to select the highest risk cells in a deterministic fashion. The second option is to use the risk analysis in a probabilistic selection of locations.
    - The number of tactics to present to officers is specified. Tactics can be turned off for a specific set of missions. If one tactic is specified, then officers are presented with only that option. Specifying higher numbers of tactics enables officers to select from several options within the application.
    - Multiple mission configurations can be active and each configuration can include one or more crime models. This means that one set of missions could focus on violence at a district level while another set of missions could address all crimes at a beat level.

Modeling Artifacts

In order to create crime forecasts, HunchLab needs to build statistical models that best represent the dynamics of each crime type. We call this process a modeling run. Modeling runs occur every few weeks to re-calibrate the model. Each modeling run has a specific unique identifier and creates several types of artifacts that are explained in this section.

Modeling Extent

This GeoTiff represents the design of the raster grid across the jurisdiction. It is created by taking a Shapefile of the areas of the county where the county police have jurisdiction (and complete data) and rasterizing it. Cell values are the percent of the cell contained within the Shapefile boundary. Any cells with values greater than 0 are included in the modeling process. Selected patrol locations (called missions) are comprised of one or more cells selected from this predefined grid.

We often specifically excluded cells that are likely to contain incorrect data or that are not useful to the operational use of the system. For example, police stations and hospitals may be removed because (1) these locations may collect police incidents with inaccurate locations due to reporting processes, and (2) these locations do not make much sense to identify as patrol locations.

Below is an example visualization of such a grid showing cells included in the analysis. In this case the county police department has jurisdiction over only a portion of the county leading.

Accuracy Files

When HunchLab builds models, it first holds back the last 90 days (duration is configurable) of crime data and builds several models that do not have access to this data. A few of the models built are baseline models that are designed to mimic what is possible for an analyst to simply do (prioritization by the number of events in a specific location or smoothed versions of these

hotspots maps along different temporal windows). The system also builds several variations of 'real' models testing two temporal weighting schemes for samples as well as different quantities of observations. All of these models are then exercised against the last 90 days of held out data with the resulting predictions assessed resulting in accuracy metrics.

We have built up a series of accuracy measures some of which we rely on heavily and some of which we used to look at but no longer do. A few of the metrics that we find useful or that warrant commentary include:

- Caught1Percent (and similar)
    - This sorts the predictions from high to low across the entire 90 day period and then measures the percent of crime captured in the top percent. On the surface this seems like a good measure, but we have moved away from focusing on it. Even if this is filtered to look at the capture rate within a particular time period (such as shift), the operational use of HunchLab does not identify the top 1% of locations across a jurisdiction. Instead, each patrol assignment area is isolated and then the best locations within that area are identified. This segmentation means that we are drawing predictions from different portions of the overall prioritization order. Because of this, we have moved toward prioritization metrics that look at overall prioritization accuracy instead of just the top end of the pririotization.
- NormalizedGini
    - This is the metric that we feel is the most important internally. The Gini coefficient can take into account non binary outcomes and so this is based upon the actual counts themselves. Since it is normalized, it also shows us what is possible to attain through a perfect model. It is difficult to explain this metric in a simple way, however, and so we rely more on AUC publicly.
    - 1 is perfect, 0 is random
- SpatialNormalizedGini
    - This metric runs the normalized gini metric for each temporal period separately and then averages across them. The goal is to identify the spatial prioritization power while removing the effect of time. This metric directly reflects the operational use of the system in that we are making selections of locations within the confines of a shift.
- TemporalNormalizedGini
    - Similarly, this measures the temporal prioritization power at a specific location removing the spatial prioritization problem from the mix. This metric does not reflect the operational use of the system, but we are using it to further determine how to increase accuracy by isolating the temporal and spatial processes.
- BootAUC.Metric (LowCI, HighCI)
    - The area under the ROC is similar to the normalized gini coefficient but on a different scale with the outcomes measured being the binary presence or absence of actual crime events. We use bootstrap sampling to build confidence

intervals around this number (95% intervals) which enables us to compare models against one another in a meaningful way. We have increasingly used this metric to explain the performance of HunchLab because there are simple explanations for the metric. For example, if the assessed AUC is 90% and if you randomly selected one example over the last 90 days where a crime occurred and one example where a crime did not occur and presented these to the model to select between, the model would correctly select the cell containing the crime 90% of the time.

  - 100% is perfect, 50% is random

The system ranks the built models along each of several of these metrics and then sums the ranks across the metrics to come up with a model score. The lowest score of 'real' models is then rebuilt including the last 90 days of data and used in production. The system is trying to select a model that performs well in both the numeric precision of the predictions as well as the prioritization power.

Model Files

Each individual model that the system builds is saved in an RData file -- a format compatible with the R language. The files for 'real' models can be quite large (up to several GBs) and can use about 10GB of RAM when loaded into R. Within the RData file are three useful artifacts:

- model.list$model - the GBM model itself built by the GBM package in R
- model.list$model.calibrator - the GAM model that wraps the GBM model's predictions built using the mgcv package
- model.list$custom.predict - a function that is used to make predictions using the two artifacts

Variable Importance Files

The variable importance CSVs are generated by the summary function in the GBM model and sum to 100 percent. (Note that the final predictions also involve a calibrating GAM model which is not reflected in these importances). Partial variable plots or the importance of variable interaction effects could be pulled from the GBM models themselves.

These files outline all of the defined variables used within a particular model. There are thematic categories of variables:

- **acs*** - These are 5 year estimates drawn from the American Community Survey data which is published at the Census Block Group level. The idea with these variables is to either (1) identify the distribution of targets of a crime such as the density of people, houses, or cars, or (2) to identify qualities of the targets that would increase their vulnerability or appeal to criminals such as household incomes or median rents, or (3) to identify environmental conditions that may make a criminal feel safer in an area such as

vacant houses, or (4) measure the ability of a community to band together to address crime issues (collective efficacy concepts)

- **row, col** - These variables represent the row and column of the location within the raster grid of the analysis. Ideally these should not be used very much and they can be a signal to us that we are missing geographic factors.
- **coverage*, density*, distance*** - These variables represent the geographic distribution of point, line, or polygon layers. The goal is to represent the built environment (places people congregate, modes of egress, etc.).
- ***prior*** - These variables represent past levels of crime along different time scales. Some of these variables are the counts at only that location (prior7) while others are smoothed versions (kdPrior7) or aggregate just the neighboring cells counts (neighborsPrior7).
- **period_since_last*** - These variables seek to represent event contagion (repeats and near repeats).
- **shift*, dow*, weekofYear, etc.** - These are calculated to represent different temporal cycles.
- **weather-*** - These variables represent weather data. Forecasts are used for future periods; actual values are used for historic periods.

Patrol Allocation Artifacts

Predictions

The prediction rasters are geographic representations of risk for a specific crime model for a specific date and shift throughout the entire jurisdiction. The jurisdiction is broken down into raster cells, and each cell is assigned a risk value -- an expected count of crime. For example, the Aggravated Assault crime model will have its own prediction raster with each cell in the jurisdiction having an expected count of events assigned to it. The prediction rasters are stored in the GeoTiff format and can be processed easily in ArcMap, QGIS, R, etc. These rasters can be combined with the event model weightings (severity weight and patrol efficacy numbers) in a raster weighted overlay operation. These weighted values are then what are used to select locations according to the mission configuration.

Missions

Missions are created by combining the prediction files with the settings in the event models and missions configurations and are stored in the HunchLab database. The default API format for missions is the GeoJSON standard but they can also be rendered as PDFs.

System Use Artifacts

API Audit Log

Every API request to the application generates an audit log record which specifies the date and time of the incoming request, the user, the nature of the request, and the type of system response. These logs are infrequently used and stored primarily for security purposes.

Application Session Records

Each time that a HunchLab user (officers, analysts, commanders, etc.) log into the application an application session record is produced. The record lasts for the duration of their login and is used to log how they intent to use the application that day (in the field or in the office). If the user is in the field, their assigned area and vehicle identifier may be logged in this record. These records are most useful for measuring user engagement with the application at a high level.

Mission Session Records

Some departments use PDFs generated from HunchLab to disseminate information to the field. If officers are directly utilizing the HunchLab interface, then mission session records are created if officers log their visits to missions within the application. Upon entering a mission, the officer specifies the tactic that they have selected from the available options. Upon exiting a mission, the officer is presented with one survey question about the mission session. For example, the officer may be asked “Did you have a positive interaction with the community?” These records also include the timestamps of entry and exit and the calculated amount of time spent in the mission and the buffer area around the mission.

Frequently Asked Questions

What differentiates you from your competitors?
Firstly, HunchLab is one of the only predictive policing software that has the ability to analyze multiple data sets to build crime forecasts. Many of our competitors only analyze past crime events, which we believe will yield less accurate results than our predictions and is more likely to perpetuate any biases present in policing. By including other data points to our models, we are more likely to build models that are based upon the root causes of crime than just its manifestation (past events). These other data sets tend to be data that is publicly available and include things such as the weather forecast for the day, the population density in an area, or the locations of schools or highway on-ramps.

The models within HunchLab have been applied to many domains with great success. In particular we are using both gradient boosting of decision trees and generalized additive models within the system. These approaches are true machine learning approaches where the model determines how interactions among variables best represent the risk of crime. Some of our competitors are rather tight-lipped about their approach in order to protect their intellectual property. In comparison, we believe that the best way to advance policing is by being transparent and using models that others can inspect and critique.

Additionally, HunchLab incorporates the concept of minimizing harm in our patrol recommendations. When making recommendations, HunchLab takes into account the amount of harm that different types of crime cause. For example, the system knows how important it is to prevent a robbery versus a burglary and makes recommendations appropriately and designs focus areas appropriately. HunchLab also is designed to take steps to mitigate the risk of harm from police presence itself. If the risk in an area is sufficiently low, it may be better for the police not to do proactive patrols and HunchLab can filter out these locations. Alternatively, if the risk is very high in an area, the natural inclination of the police may be to saturate the area with patrol resources, but very high degrees of police presence may create harm by creating a sense of harassment by the police. Just like a doctor will prescribe the appropriate dose of medication to optimally treat an ailment -- too much medicine can be quite harmful -- HunchLab meters out the proactive patrol dose in proportion to the risk in a location in an intelligent manner.

Given the inherent biases in policing data, won't the predictions themselves also be reflective of the bias?

Crime data suffers from several forces that introduce biases to varying degrees. Any model making predictions based upon these outcomes will naturally reflect such biases. No data is perfectly accurate and by its definition, a statistical model is also an imperfect representation of reality. This question often arises due to the disparate impact of policing and incarceration on communities of color, so our answer will cast this issue in that light.

There are a few sources of bias within policing data. One source of bias is reporting rates. Low income and minority individuals are less likely to report crimes to the police[12]. These lower reporting rates may be due to fear of deportation, distrust for the police, or the simple lack of necessity for a police report. For instance, if your house is burglarized, it is quite unlikely that the police will recover your possessions, but a police report enables you to file an insurance claim. If you do not have insurance because you are low-income, you have little reason to involve the police. In time, it may be possible to account for reporting bias by adjusting crime data to reflect a corrected level of activity based upon separate studies of reporting inequities. For now,

---

[1] Hindelang MJ, Hirschi T, Weis JG. Measuring delinquency. Beverly Hills, CA: Sage; 1981.
[2] Press, Eyal. "Do Immigrants Make Us Safer?" *The New York Times*. The New York Times, 02 Dec. 2006. Web. 11 Nov. 2015.

however, we have not done work to address this bias within HunchLab because correcting for it will only amplify the focus on low income and minority communities.

A second source of bias is enforcement bias. This type of bias varies across crime types. This bias is less present in major crimes such as homicides, robberies, assaults, or burglaries than in crimes such as drug-related or nuisance crimes. Crimes largely fall into one of three categories: (1) crimes that the public reports such as major incidents, (2) crimes that the public reports which don't immediately lead to an incident report such as noise complaints or drug activities, and (3) crimes that the public does not report but are instead initiated by the police. We focus our models to address this issue. In particular, we will use police report records to model crimes where an enforcement bias is less of an issue such as major events. For quality of life type crimes, we tend to use records that reflect the public's call for services, which does not suffer from an enforcement bias.

Given the limitations in crime data, a fundamental question is whether it is better for the police to engage in location-specific policing activity given such limitations or for the police to not conduct location-specific policing activity. This question often arises due to the disparate impact of policing and incarceration on communities of color, but when modeling crime geographically, it is important to realize what is being modeled.

The location of a crime event represents the location of the victim of the crime (whether a person or object). These victim locations are based upon the places where people live, work, go to school, and socialize or on the journeys between these places. Offenders often do not commit crimes exactly where they live but rather travel short distances to commit crimes[3][4]. If crime is concentrated in communities of color, it suggests that those communities are disparately victims of crime. Most people would agree that violence disportionately affects communities of color in the US. By preventing a crime from occurring, we prevent people from becoming victims. By preventing a crime, however, we also prevent an offender from being charged with the crime, which can reduce incarceration. We believe that a geographically focused approach with meaningful police tactics that build relationships with the community makes sense.

Is this like minority report?

No this is not like Minority Report. The premise of Minority Report was that people were arrested for crimes they had not yet committed. We are forecasting risky locations for crimes to occur, with the goal of no one being arrested because the crime is prevented.

---

[3] Brantingham, P.J. and Brantingham, P.L. (1978). A Theoretical Model of Crime Site Selection. In Crime, Law and Sanctions (M. Krohn and R. Akers, eds). Sage; Beverly Hills.

[4] Brantingham, P.J. and Brantingham, P.L. (1981). Environmental Criminology. Sage; Beverly Hills.

Will crime go down if departments use your tool?
It depends. We do believe that our tool provides accurate forecasts of risky areas for crime. However, that is not enough to garner a reduction of crime. Firstly, the officers that have been assigned to these mission areas have to follow the recommendations and patrol those areas. Secondly, it also depends on what sort of tactics that the officers are using in these risky areas. We believe that just the presence of officers in areas are enough to deter crime, but there are likely to be more effective tactics that could be implemented. A reduction in crime can only be attributed to the combination of the factors at play. HunchLab can be a part of that but it is a gross misrepresentation to attribute a crime reduction to just a specific analytic process.

Finally, as every department is different and the problems that they face are different, some departments may have different levels of success when using HunchLab. So far in the departments that have used us long enough, we have seen a decrease in crime, but that could be different for other departments. We have a low-cost pilot program so that a department can implement HunchLab operationally and see how well it works for them before committing to a long-term contract.

Does this change anything with how the police will interact with the community?
At its core, HunchLab aims to find better ways for police to allocate resources more efficiently. Our hope is that, in directing officers to where their attention would be most effective at deterring crime, the incidence of crime will fall. Moreover, by giving officers more context about the areas that they're going to, we hope that it helps them interact with the community in a more cognizant way. If they know that a particular area that they're visiting is at risky for aggravated assault, they now know to be more sensitive to these issues when patrolling the community.

We do not see any immediate impacts that may change how the police the will interact with the community in a negative way. Moreover, HunchLab may bring attention to new places that officers have not thought to patrol before. Because HunchLab also recommends the amount of time an officer should patrol in certain areas and because Advisor also recommends effective tactics, we believe that'll we'll improve the effectiveness of policing.

Has HunchLab or predictive policing been tested?
Greensboro PD in North Carolina has done a study, testing the accuracy of HunchLab's predictions as well as seeing how it fits in with their department's current practices. Overall, they found HunchLab to be more accurate than their current tools of analysis, and as such helped reduce crime rates overall. They also found that using HunchLab didn't prove disruptive to their current departments practices, integrating pretty seamlessly.

The Philadelphia Police Department conducted a randomized controlled experiment that used HunchLab to evaluate how heavy police saturation of different types of units affected crime rates. The results of this study found some statistically significant effects.

Overall we think the question of “does predictive policing work” is the wrong question as it is too generic. Our system’s identified locations could be used to conduct various types of interventions. Some of those interventions likely prevent crimes; some of those interventions may improve community relations; and some of those interventions may worsen community relations. In other words, the real question to ask is what is being done in the field and what are the effects both positive and negative. This question has less to do with how the locations are selected and more to do with the interventions..

When was the company founded?

Azavea has been around for about 15 years. In addition to HunchLab, we provide other products as well as professional services for geospatial analysis with a focus on civic applications. As a registered B-corp, we focus on projects that have a civic and social impact on society.