Final Capstone Proposal

**D**ynamic **I**ncident **R**esponse **E**nvironment

# Driving issue:

Whether you fell off the roof, have severe abdominal pain, are about to have a baby or your heart has stopped beating the expectation is calling 911 will get someone to come to your location who is properly trained as quick as possible. There many factors that influence how quickly the emergency responders will arrive. Both the individuals involved in the incident and the crews responding have a sense of urgency associated with the incident and often the outcome of the more serious incidents can be tied to the amount of time it takes for crews to arrive on scene.

There are two primary factors relating to emergency response times. The first factor is the amount of time the dispatcher takes to collect your information and determine exactly what has happened and who they need to send to help you. The second factor is where the emergency crew is responding from in relation to the incident location.

The first factor can be minimized with what is called pre-dispatching, allowing the dispatcher to start sending crews prior to all the information being collected on the more serious incidents. Typically, their pre-dispatch questions are asked immediately after collecting a phone number and address. These questions are related to caller/patient safety, breathing, pulse, bleeding and the presence of weapons.

The second factor is really a combination of how many active incidents are already going on, reducing the number of crews available to respond, and the location of the available crews in proximity to the incident in question. In many systems there are predesignated locations around the city and a crew will be assigned to that location either for their entire shift or short periods of time determined by the current number of available crews. How the crews are posted is also determined by emergency response guidelines which can specify what the maximum response time is for different regions in the response area. These guidelines in combination with previous incident data are extremely helpful for preplanning the posting locations.

The guidelines and past data do not address the issue of actually assigning the posts so that the crews are in the optimal position for calls in the near future. Fire Departments and Ambulance providers already have begun doing forecasting to allow them to predict how many ambulances will be needed to meet the incident volumes over the course of a week. Some organizations have also taken it a step further and have identified trends in call location that correspond to higher call volume areas and weight their model to staff more ambulances in higher call volume areas during peak hours. However, it is still up to the dispatcher to determine where to move the crews to in response to new incidents and where to assign crews as they become available to respond again.

# Resolution:

My goal is to optimizing our second factor of **crew location** is to develop a more dynamic and adaptive model for identifying where to put the crews. To do this my model will utilize multiple years of past data to define ideal posts (cluster centers). The model will break down time into either 30 minute or 1-hour bins of time, depending on model accuracy and efficiency, to allow for rapid data processing and not irritate crews by having them move too often. The model will incorporate recent (two weeks and the time bins from a month prior and a year prior) and the previous time periods (last twelve-time bins). I also plan on incorporating a feature that calculates the error from the last 24 hours’ worth of predictions. All of this will be used to predict the likelihood of where incidents will occur for the current and next two time bins. This will result in reduced response times and keep staffing to a minimum while meeting any guidelines set by governing bodies.

I anticipate our solution will utilize data cleaning, normalization, feature creation and importance selection. The first major step will be defining the number and location of the cluster (posting areas) that will be used. I intend to implement clustering through some form of kmeans. Our next major step will be creating manageable time bins based on the response data and department specifications. The last step will be implementing a predictive model to not only predict how many incidents will occur in the next time period (bin) but the next 2-5 time periods and where they are the most likely to occur within the response area (clusters) For the prediction and error reduction portion of the model I will likely utilize ARIMA, random forest and XG Boost, SGD, a simple moving averages and potentially a combination of models to attach the number and location of incidents

# Expected hurdles:

The biggest challenge I will face is the complexity of the incident volume, location and severity as the time of day, day of the week and day of the year affect the data. I will also run into issues with data entry errors around incorrect addressing of incidents, and time stamps that are not formatted appropriately. I may also run into a need to change the duration of a time bin depending on the calculated average call duration (called into dispatch to crews being available to respond to a new call). I can already anticipate there may be some difficulty combining the location and the number of incidents into a prediction. I also expect that I will need to address how to deal with regions in the response area that typically have very low incidents per day and periodically have periods where they see a drastic increase in call relating to special events (concerts/festivals/events, transportation related emergencies, weather/environment related incidents, illicit drug activity, seasonal activities).

End Product:

Rollout:

The final product is conceptualized to allow for historical data to be uploaded into the data frame as part of the training phase. We will also set parameters through an initial “calibration” for local guidelines on response times and staffing levels. We will also need to define the geographical boundaries and potentially add geographical markers to our mapping visuals.

The actual user is intended to be a dispatcher, supervisor and potentially some of the responding units. There will also be a pipeline for live data coming in from the dispatch server and services like the weather and traffic feeds to allow for accurate predictions. I envision the system creating a notification of a new incident and then the user utilizing the model to produce the ranked locations for where to send the next available crews to stage while they wait for the next anticipated incidents. I also envision there will be visuals for 24-hour incident volumes and a heat map of the recent trends in call location along with a hot list of the last two time periods worth of incidents in a table.

Implementation:

I will be using the data from the City of Cincinnati to build our product. We will primarily utilize the longitude, latitude and incident creation time provided. I Will create new features for year, month, day, day of the week, hour and minute that an incident was created. We will also create cluster centers and time bins as additional features. Once the model is up and running, we could add in weather, traffic, call type, weighted average incidents per cluster per time bin for the previous day, two weeks, one month to date and time and one year to date and time.

Resources:

Incident/Response Data:

Cincinnati Fire Incident Data

<https://data.cincinnati-oh.gov/Safer-Streets/Cincinnati-Fire-Incidents-CAD-including-EMS-ALS-BL/vnsz-a3wp/data>

The site provides basic analytics about the data and allows for filtering. The data goes back to 2015 and include response location, time the incident was created (someone called 911) times for when the dispatcher notified crew, when the first crews arrived and when the crews have cleared that incident. There is also data on the type of call and additional information about the location (assigned station, neighborhoods etc.…)

Traffic data for inclusion future

Cincinnati Traffic Statistics:

TomTom

<https://www.tomtom.com/en_gb/traffic-index/cincinnati-traffic#statistics>

Cincinnati Weather data

Current:

<https://www.accuweather.com/en/us/cincinnati-oh/45229/weather-forecast/350126>

proposed weather API tie in for product

<https://www.athenium.com/products/api/>