Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

RIDE ALONG: USING SIMULATION TO STAFF THE ITHACA POLICE DEPARTMENT

Christopher Archer
Matthew Ziron
Jamol Pender

Operations Research and Information Engineering
Cornell University
Ithaca, NY 14853, USA

ABSTRACT

In this work, we combine data science, queueing theory and stochastic simulation into an analytical framework to examine the operations of the Ithaca Police Department (IPD) and develop novel solutions that address recent staffing shortages of the IPD. With 3 years of IPD call data, we construct and simulate a novel multiple-dispatch queueing process for modeling the IPD's operations. Unfortunately, using the data alone does not capture the real dynamics of IPD's operations, which made it necessary to augment the call data with qualitative assessments from police ride-alongs. Our analysis augmented with ride-along assessments accurately replicates metrics such as emergency response time and officer utilization. Although it is difficult to fully capture the claimed strains on the department through data alone, hiring additional officers should make an observable improvement in the performance of the IPD.

1 INTRODUCTION

Ithaca is a mid-sized college city in Tompkins County, upstate New York. It covers approximately 6 square miles with a permanent population of 31,000, a Cornell University student population of 24,000, and an Ithaca College student population of 6,000. The Ithaca Police Department (IPD) exercises jurisdiction over much of the city, excluding the two universities and various villages inside the town.

The Ithaca Police Department is in charge of maintaining the safety of the City of Ithaca and enhancing the quality of life of its residents. Recently, due to many officers retiring or being unable to work, there has been a shortage in the number of active police officers (https://ithacavoice.com/2018/10/police-staffing-still-in-question-after-common-council-budget-meeting/). This on top of a lack of approval for additional funding and an overall decreasing applicant pool has led to many problems in the staffing and operations of the IPD (https://www.ithaca.com/news/ithaca/ipd-seeks-funding-for-additional-officers-yet-again/article_9d47648c-d30d-11e8-b9aa-27cad6619b1d.html). For example, the contractual minimum number of officers on patrol during any one 8 hour shift is 6, and often times the IPD has only been able to have 4 officers on patrol per shift. Officers have reported feeling overworked, stressed, and incentivized to take a large amount of overtime, which many officers comment is unsustainable for the long term.

As a result, we were approached by Mr. Dan Cogan, who is the City of Ithaca Chief of Staff, and Sergeant Loretta Tomberelli to understand whether the IPD needed to hire additional officers by using IPD call data. The data we were given was in the form of *cards*, which are brief reports which summarize the important aspects of the calls that are sent to the Dispatch center. Information such as time of dispatch, officers involved, location, and time of call completion are all included in these cards. However, despite having very detailed data, we found that these cards or data do not give a complete picture of all of the work that police officers do. For example, we observed that a significant amount of time is spent doing followup work, such as writing summary reports of incidents and conducting investigations. These activities

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

are generally not included in the call data, so we find that a purely data-driven approach is not effective in actually modeling police dynamics. To supplement the data, we also went on ride-alongs with patrol officers Mr. Michael Meskill and Mr. Kurt Soderholm to gain more insight into how an officer's time is occupied during a shift. To understand the impact of the information that is not collected from the data, we investigated the difference between this purely data-driven ”consumer perspective” modeling and the more holistic ”officer perspective” approach later in the paper.

1.1 Goals of Project/Analysis

Broadly, the goal of our project is to use the data from the Ithaca Police Department's past calls to see if there are any operational changes that can be made within the department which can decrease the stress/workload on the officers. We tackle this problem by leveraging the IPD call data and ride-along information to build a stochastic simulation for IPD's call process from start to finish. We found that we could model the IPD's call process as a non-standard multi-server queue. It is non-standard because multiple agents may be needed for each individual service request. For example, a report of gunfire would certainly call for multiple officers, while a noise complaint might only require one officer. By building a stochastic simulation model of patrol cars responding to calls in Ithaca, we can estimate important IPD performance measures like emergency-call response time and officer utilization. Moreover, the stochastic simulation driven by our data analysis will enable us to study the impact of different staffing levels on these performance measures.

2 POLICE DATA

In this section, we describe the data we collected and analyzed to give readers a sense of the large amount of data cleaning and analysis that was necessary for building our models.

2.1 Explanation of Data

We received data from IPD Sergeant Loretta Tomberelli. We were given the data in the form of *cards*. A card represents a specific incident the police department handles. The format of these cards contains several important details about the call which will be vital for our analysis, namely:

- *Call ID*: an arbitrary ID that is unique to each call
- *Dispatch*: The time in which a 911 call is received
- *Enroute*: The time when an officer chooses to respond to a call
- *Completed*: Time when an officer has completed their initial assessment of the call
- *Crime*: Label for what kind of crime is being reported
- *Responding Officers*: The list of officers responding to a crime

The general process of a card being logged is as follows. A call for a crime is received at the **Dispatch time**, the first officer choosing to take this call is the **Enroute time**, and when the first officer finishes the initial assessment they log their **Completed time**. It is often the case that multiple officers are called to handle a single call, but the times shown in the data only correspond to the first officer who responds. We show a sample path of a call into IPD in Figure 1 where the red arrows correspond to actual data points that were were able to collect from the cards.

2.2 Issues with the Data

Upon receiving the data, we found that the data was not clean and suffered from many problems. The three main problems with the data were: (1) the data were formatted inconsistently (2) many data points were missing and (3) there were many spelling errors and inconsistencies. To get a better sense of what the

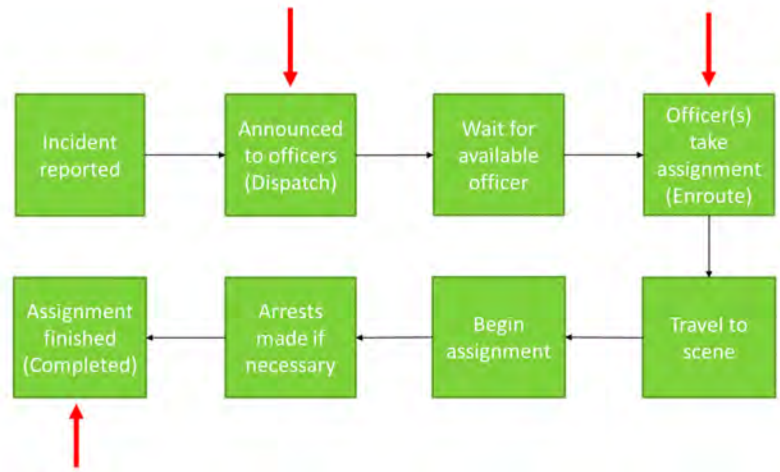

Figure 1: The progression of a general 911 call, with red arrows pointing to the points in time present in the data.

original (uncleaned) looks like, we provide a snapshot in Figure 2. In Figure 3 we also show a snapshot of the cleaned data, which is much easier to analyze from an analytics perspective.

1	Number Time and Date Nature Address Loctn Dsp
2	--------- ---------------- -------------- ------------------------- ----- ---
3	Agency: IP Zone:
4	Responsible Officer: M ORSAIO
5	I17-01765 00:52:01 02/01/17 WELFARE CHE CK 604 S ALBANY
6	Responding Officer Dispatched Enroute
7	----------------- ---------------- ---------------
8	M ORSAIO 0:52:01 1:07:40
9	L MORSE 0:52:01 1:07:40
10	G DUPAY 0:52:01 1:07:40
11	M COWEN 0:52:01 1:07:40
12	Responsible Officer: G DUPAY
13	I17-01766 01:48:02 02/01/17 TRAFFIC OFFENSE W STATE ST / N MEADOW ST,
14	Responding Officer Dispatched Enroute Completed Total
15	----------------- ---------------- -------------- ------------ -------- G DUPAY 01:48:02 N/A 01:51:21 N/A
16	L MORSE 01:48:02 N/A 01:51:21 N/A
17	Responsible Officer: M ORSAIO
18	I17-01767 04:52:47 02/01/17 ALARM POLICE 160 COMMERCIAL AVE; SOLID
19	Responding Officer Dispatched Enroute Completed Total
20	----------------- ---------------- -------------- ------------ -------- M ORSAIO 04:52:47 04:55:49 04:59:55 0:4:6
21	L MORSE 04:52:47 04:55:49 04:59:55 0:4:6
22	Responsible Officer: J PAPKOV

Figure 2: Original Data

Since the IPD call data is held in a central repository, Sergeant Tomberelli was able to send us one CSV for each month of years 2016, 2017, and 2018. However, we found that the data collection process was not uniform. In fact, we observed that each month could be in one of many different data formats and there was no clear information that would allow us to know which format it would be. Thus, the decision was made to use the most convenient and accurate format and remove the rest. The months of data that were ultimately used in this paper were

1. January, February, September, November, and December of 2016.
2. January, February, March, April, August, September, and October of 2017.
3. February, March, and August of 2018.

In these 15 months of data, we observed that almost 30,000 incidents were recorded yielding about 2,000 incidents per month. The data size was large enough to work with, but we note that much of the summer

	A	B	C	D	E	F	G	H	I	J	K	L
1		Address	Completed	Date	Dispatch	Enroute	ID	Offense	Officer1	Officer10	Officer2	Officer3
2	0	HECTOR ST	########	4/1/2017	########	########	I17-05208	TRAFFICOF	J BALDESSARE		L MORSE	
3	1	116 E STAT	########	4/1/2017	########		I17-05209	LOCALLAW	J JIMENEZ IP		S GARIN	
4	2	E BUFFALO	########	4/1/2017	########	########	I17-05210	TRAFFICOF	J JIMENEZ IP		L MORSE	E DOANE
5	3	W BUFFAL	########	4/1/2017	########	########	I17-05211	TRAFFICOF	L MORSE		J BALDESS	S GARIN
6	4	540 W STA	########	4/1/2017	########	########	I17-05212	ASSIST	J BALDESSARE		J JOLY	J PAKKALA
7	5	E BUFFALO	########	4/1/2017	########		I17-05213	DISORDERI	E DOANE			
8	6	610 CHEST	########	4/1/2017	########	########	I17-05214	NOISECMP	J BALDESSARE		E DOANE	
9	7	108 PENNS	########	4/1/2017	########	########	I17-05215	BURGLARY	K FIELD		J PAKKALA	E DOANE
10	8	837 N AUR	########	4/1/2017	########	########	I17-05216	ASSIST	P KIMMICH		C VANCLEI	M GRAY
11	9	616 W CLII	########	4/1/2017	########	########	I17-05217	DOMESTIC	J ARSENAULT		N LOPEZ IF	ZZ S HOFFI
12	10	314 W STA	########	4/1/2017	########	########	I17-05218	ALARMPOI	C VANCLEEF		P KIMMICI	M GRAY
13	11	120 E GREE	########	4/1/2017	########	########	I17-05219	REPOSSESS	S CREWS			
14	12	RT 13 / DEY	########	4/1/2017	########		I17-05220	TRAFFICOF	S CREWS			
15	13	602 TOWE	########	4/1/2017	########		I17-05221	SPECIALDE	P KIMMICH		M GRAY	J BUFFONE
16	14	801 TAUGH	########	4/1/2017	########	########	I17-05222	PARKINGPI	N LOPEZ IP		C VANCLEEF	
17	15	610 W STA	########	4/1/2017	########		I17-05223	ASSIST	C VANCLEEF		N LOPEZ IP	
18	16	12 YARDLE	########	4/1/2017	########	########	I17-05224	ASSIST	J ARSENAULT			
19	17	219 CLEVE	########	4/1/2017	########	########	I17-05225	SUSPICIOU	T CONDZELLA		P KIMMICI	J NELSON
20	18	748 S MEA	########	4/1/2017	########	########	I17-05226	PDACCIDEI	J COLE		R DUPAY I	J BYRD
21	19	100 BLOCK	########	4/1/2017	########		I17-05227	TRAFFICOF	R DUPAY IP			

Figure 3: Cleaned Data (All cells with ### encode date-times)

months are missing. However, summer months are typically less busy while students are away from school. We also should highlight that the data corresponds to the time frame where the IPD has felt that their workload was too high and the IPD made formal requests to the Mayor's office in Ithaca to hire 20 additional officers (https://cnycentral.com/news/local/ithaca-police-department-mayor-dont-see-eye-to-eye-on-staffing).

This cleaned data allowed for easier analysis since we now had a consistent indexing of rows, something notably absent in the old data. Additionally, the data also had several missing entries, the majority of which were Enroute times. We learned this was mainly due to officers being in close proximity to a call, as officers who are already near a crime do not generally log the time they respond to it. We chose not to impute these data, and will discuss the effects of that in the next section.

3 EXPLORATORY DATA ANALYSIS OF ITHACA POLICE DEPARTMENT CALL DATA

In this section, we conduct an exploratory data analysis of the IPD call data. This data analysis is instrumental to our paper as it serves as it gives us a sense of how to construct the simulation for various call types. Moreover, we are able to get accurate estimates of call types, number of officers involved, duration of the incident, etc.

3.1 Ithaca Police Department Call Data

From the analysis of past call data, we found several important pieces of information relating to how calls are received and processed within the police department. The first of which being the distribution of inter-arrival times (the time between one emergency call and the next) was found to be almost perfectly exponential, see for example Figure 4. This observation is important because the exponential distribution is the building block of the Poisson process and the Poisson process is well understood in the stochastic simulation community. Thus, our observation of the inter-arrival times being almost exponential allows us simulate the calls to the IPD using a Poisson process, which reflects the statistical independence of each call. Thus, we do not need to resort to more complicated arrival processes like Hawkes processes i.e. (Daw and Pender 2018a) and (Daw and Pender 2018b).

In Figure 4, we also observe that the mean inter-arrival time of each call is 21.67 minutes and the median inter-arrival call time is 12.4 minutes. This implies that the police department receives an emergency call approximately once every 21.67 minutes. It also hints that the inter-arrival call data is positively skewed as well. The inter-arrival time focuses primarily on the *Dispatch time*, however, it does not reveal any information about the actual progress of the call. Thus, we also analyzed the distribution of the time between a call's arrival and an officer traveling to address the issue (the Enroute - Dispatch time). On the right of Figure 4, we display the distribution of the gap between Enroute - Dispatch. Similar to what we

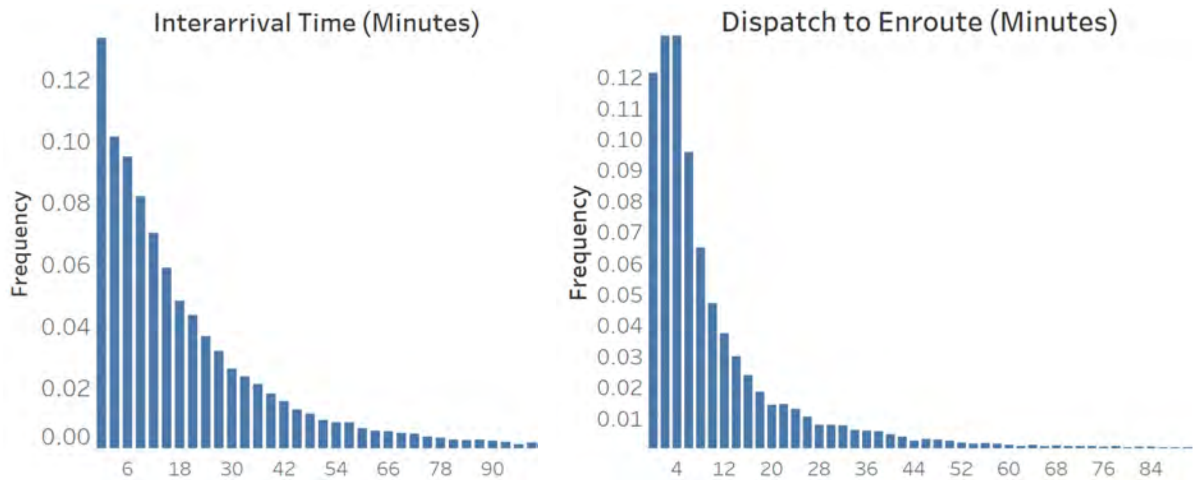

Figure 4: Distribution of Inter-arrival times (mean = 21.67 min, median = 12.4 min) (Left). Differences in times between Dispatch and Enroute (mean = 10.5 minutes, median = 5.2 min) (Right).

observed with the inter-arrival times, the distribution of the gap between Enroute and Dispatch appears to be approximately exponential (although a little less so than the inter-arrival times). We also observe that the mean gap is approximately 10.5 minutes and the median gap is 5.2 minutes, which also implies the Enroute - Dispatch gap is also positively skewed.

We also analyzed the time it takes for a call to get completed (Completed - Enroute) and the total time a caller is being serviced (Completed - Dispatch), shown below in Figure 5 on the left and right respectively. We observe in Figure 5 that the Completed - Enroute gap is approximately exponential as well. The mean gap is 16.5 minutes and the median gap is 10.3 minutes. In the case of the Completed - Dispatch gap, once again we see the right plot of Figure 5 that the distribution is approximately exponential. The mean is roughly 25.3 minutes and the median is 17 minutes.

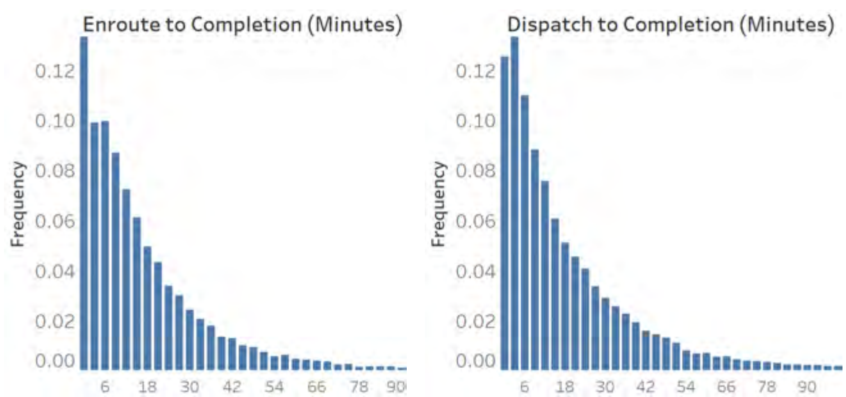

Figure 5: Differences between Enroute and Completed (mean = 16.5 minutes, median = 10.3 min) (Left). Differences between Dispatch and Completed (mean = 25.3 minutes, median = 17 min) (Right)

Now that we have a good understanding of the time gaps between events, it is also important to understand the resources needed per call. Unlike many queueing models that assume that there is one server assigned to each request, this is not the case in emergency services like the IPD. In Figure 6, we calculate the number of officers needed per call. We find that 55% of the time only one officer is needed, 27% of the time 2 officers are needed, 10 % of the time 3 officers are needed, 4% of the time 4 officers

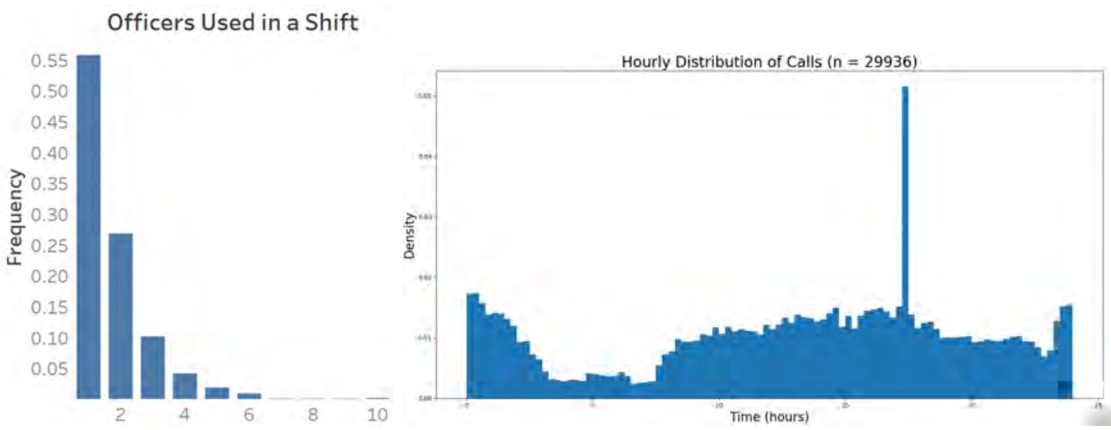

Figure 6: Distribution of Officers Needed per Call (mean = 1.8, median = 1.0) (Left). Distribution of Calls over an Average Day (bin size = 15 mins) (Right).

are needed and the maximum officers need was 10 during the 3 year time period. We find that on average 1.8 officers are needed per call and the median is 1 officer per call.

3.2 Understanding the Top Offenses

Now that we have an understanding of how many calls are coming in and how many resources they require, we now can begin to analyze the data in a more refined way. To this end, an important factor to consider beyond how often calls come in and are processed is *when* calls come in and *what* type they are. Since every shift of the IPD covers a different part of the day, knowing the time of day when calls are most frequent is very useful for distributing staffing levels between the three shifts. We found the 10 types of calls the IPD receives most frequently to be: ASSIST, NOISECMPLNT, PARKINGPROBLEM, PDACCIDENT, PROPERTYCMPLNT, SUSPICIOUS, THEFT, TRAFFICCMPLNT, TRAFFICOFFENSE, WELFARECHECK. We chose to use these offenses as the basis for our simulation because they comprise about 80% of all calls in the data. Figure 7 gives the top five call types which comprise about 55% of all call volume.

In addition to breaking the data down by time and day, it is also natural to wonder if some of the important trends hold when observing particular offenses. To explore this, we analyzed the top five offenses: ASSIST, PARKINGPROBLEM, PDACCIDENT, PROPERTYCMPLNT, and TRAFFICOFFENSE in Figure 7. In particular, we look at the inter-arrival time, dispatch to completed time (the total time a caller spends in the police system), and the number of officers used per call. One observation is that ASSIST and PROPERTYCMPLNT experience the 5:15 PM spike while the other offenses do not. Additionally, PARKINGPROBLEM peaks in the morning as people wake up and notice parking problems (e.g. they cannot get out their driveway because someone has parked in front of it). Moreover, the exponential distribution of the gaps between aggregated calls and their progress is also present for sufficiently common offenses. There are some notable exceptions, such as how dispatch to completed times are approximately gamma distributed for PDACCIDENT, but the property generally holds.

Some of the rarer calls, such as stabbings and assaults (which occur 6 and 55 times or .0224% and .205% of the time, respectively) take up resources disproportionate to their commonality, but there are special measures for when especially serious crime occurs so that normal operations can continue, such as receiving aid from the county sheriff or from other towns. So it may be of note that looking at a small number of common offenses can still give a reasonable picture of what an "average" day might look like.

For example, this information implies that if one of the more common offenses is especially costly in man-hours to deal with directly, it may be a non-obvious financial problem as well as a societal one. We are unable to explore the many opportunity costs such as paperwork time and following up with citizens, which

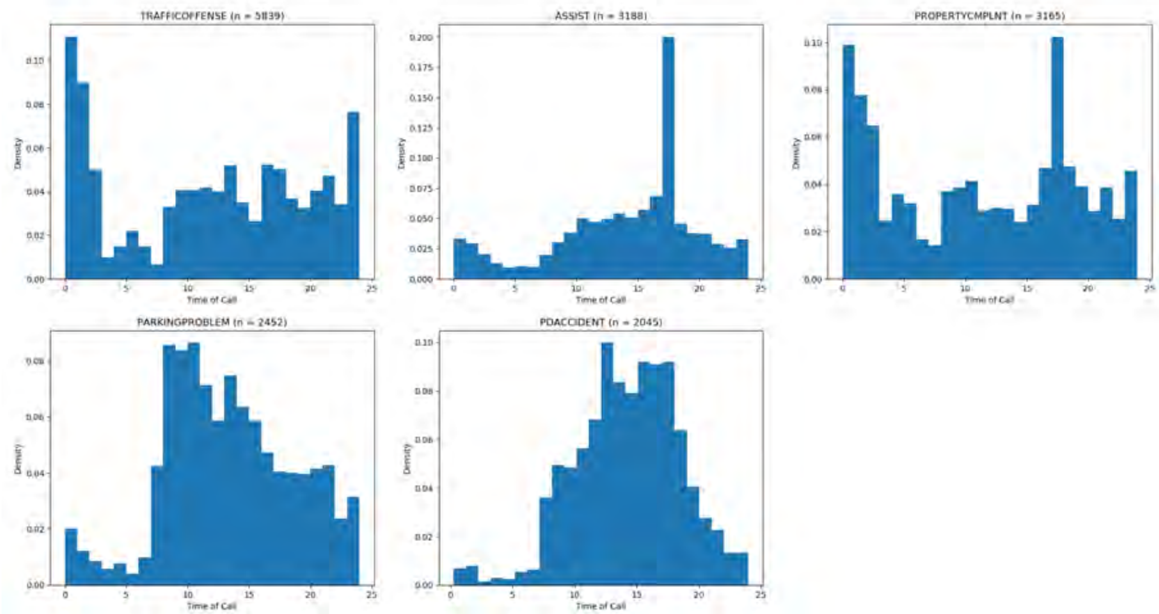

Figure 7: Hourly Distribution of Top 5 Offenses

are substantial, but we can calculate some opportunity cost which affects both the citizens affected and the police department by multiplying average officers responding by the time from dispatch to completion and the number of times that offense is reported in an hour. Calculating this measure of man-hours per hour for the top five offenses, we obtain the following values from greatest to least, 3.2 (PDACCIDENT), 1.6 (PARKINGPROBLEM), .70 (TRAFFICOFFENSE), .60 (ASSIST), .48 (PROPERTYCHECK).

3.3 Ithaca Police Department Shift Data

To supplement the call data, we were also given data which illustrated how many officers were working every day in the years of 1999, 2004, 2009, and 2014. Although not necessarily displayed here, the data revealed an important trend: the size of the patrol division has been largely constant over the years we have data on, but the number of people working overall on a given shift has decreased substantially.

Figure 8 shows the shift staffing for any one day, of which data is available for all months of 1999, 2004, 2009, and 2014. In this data, **Shift 1** refers to the midnight shift (11 pm day before - 7 am), **Shift 2** refers to the day shift (7 am - 3 pm), and **Shift 3** refers to the afternoon shift (3 pm - 11 pm). The "Patrol" data is key here, because that is the division of the IPD that is primarily responsible for responding to citizen calls. There are many other duties officers can take on during their shift, and those were counted as well. On the left of Figure 8 we plot the Patrol staffing via a box plot and on the right of Figure 8 we plot the Total staffing via a box plot. The data supports claims by the police department that staffing in the Patrol division has been maintained at the expense of other departments, namely in shift one. In fact, the median number of officers working has decreased from about 25 in 1999 to about 15 in 2014. In combination with the steady staffing in patrol, this suggests that a substantial part of the drop in service may not be in response times, but in how long larger scale tasks and investigations may take, and in non time related metrics such as number of traffic tickets issued, increased overtime pay, and number of minor calls not attended to. It is also true that crime in Ithaca has decreased over this time period. It is worth noting the presence of zeros in the charts of numbers of officers on patrol. It is expected that this is due to errors in entering the data - that some days may have been skipped and have zeros populating the empty fields. This said, the data comes directly from the IPD and should therefore be considered reliable.

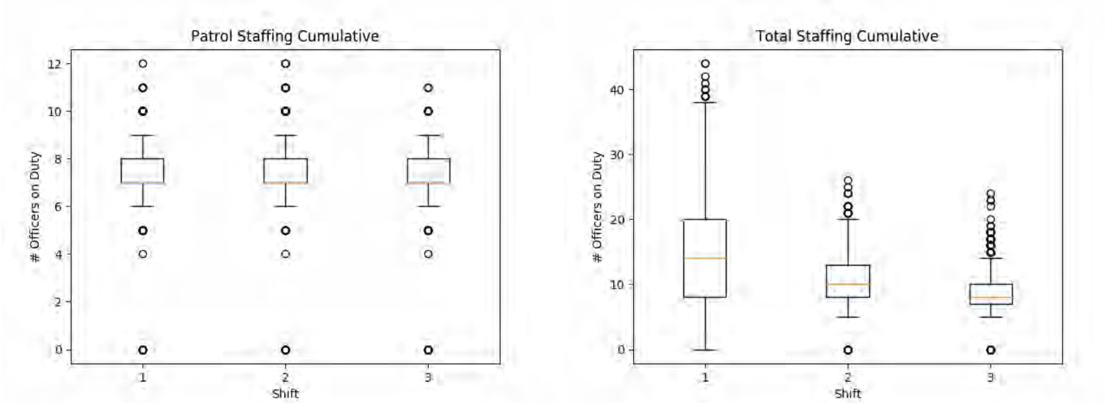

Figure 8: Distribution of Officers on Patrol (Left) and Duty (Right) Broken Down by Shift

4 SIMULATING THE ITHACA POLICE DEPARTMENT

In this section, we describe how we simulated the IPD's operations. We chose to model the IPD's operations through the lens of queueing theory. Queues have been used for a very long time to model various service systems in telecommunications, entertainment, transportation, healthcare and even autonomous vehicles, see for example (Brown et al. 2005), (Zeltyn and Mandelbaum 2005), (Novitzky et al. 2019), (Massey and Pender 2018), (Nirenberg et al. 2018), (Palomo et al. 2020), (Kahara and Pender 2017), (Pender et al. 2020), (Daw et al. 2019), (Hampshire et al. 2020). The model we use in this paper is a Markovian multi-server queue where the number of servers requested per call is random. In this model, we must specify how it operates when there are not enough servers to initially accommodate a request for help. In the situation where no police officers are available, we have the call wait until one police officer is available and that officer is sent to help that issue and no other officers are sent. In the situation where the call requests more officers than are available we send the idle officers to the scene and no other officers are sent. This is clearly an approximation to the realistic scenario, but still offers insight into IPD's process by which they handle calls.

In most queueing systems, the number of servers requested is equal to one. However, as noted earlier the IPD can send multiple police officers to handle a call, and sometimes this is required for high priority calls like shootings and break-ins. In this system, we assume that calls into the police station (or 'customers' who dial 911) arrive according to a Poisson process with a fixed rate λ (0.05 min^{-1}). Moreover, each arrival is attached to a request for a random number of officers in order to service the call. From the IPD data, we observe that it is common for 1 or 2 officers to report to the scene of a call, but there are some situations (although with lower probability) where more officers are needed. This specific kind of queueing model where the number of servers is random is not very well studied in literature, however, there are a few papers that attempt to do so, see for example (Brill and Green 1984), (Federgruen and Green 1984), (gre), (Green 1981), (Green 1984), (Schaack and Larson 1989), (Verma et al. 2010), (gre), (Ignall et al. 1974). Our application to police settings and the performance measures we study are novel in this random number of servers setting.

The benefit of this simulation is that it takes in the information that we are able to measure from IPD call data and turns it into insightful information like waiting times or response times for calls. In our simulation, we input the average time between calls, average time to complete a call, and average number of officers per call, and then measure how long people calling 911 spend waiting, what percentage of the shift an officer spends answering calls, the number of people being served in total, and the number of officers in use at randomly selected points in time. The outputs of the simulation can be used to understand the caller perspective and give Tompkins county a complete IPD system perspective. Before we went on our ride-alongs, we assumed that the time to complete a call could be approximated by what was measured from the data. This gave an impression that a majority of calls finish in a few minutes, with only a few taking substantially longer, with an overall average service time of about 20 minutes. We also did not know how many officers were on duty during a normal shift, so we assumed they had 7 officers per shift as per contractual minimum. However, the simulation allows us to obtain insight about how increasing or decreasing the number of officers will impact the community.

4.1 SIMULATION RESULTS: DATA ONLY

First, we detail the results of our naïve, pre-ride along approach. Each simulation was run with either 5,6,7, or 8 officers in a shift, an average inter-arrival time of 20 minutes, and an average service time of 25 minutes as given by the call data. We also assumed that the number of officers needed was given by the call data distribution given in Figure 6. However, we truncated the distribution at 6 officers and re-normalized the probabilities since needing more than 6 officers is very rare. Thus, the probabilities that were used in the simulation where $p_1 = .56, p_2 = .27, p_3 = .10, p_4 = .04, p_5 = .02, p_6 = .01$.

Table 1: Markovian Random Server Model (6 Officers).

	Pre Ride-Along	Post Ride-Along	Difference
Mean Waiting Time	.0537	3.3461	3.2924
Mean Queue Length	.2446	.5709	.3263
Mean Number of Officers Used	.0146	.2915	.2769
Mean Number in System	.7504	1.9194	1.169

Table 2: Markovian Random Server Model (7 Officers).

	Pre Ride-Along	Post Ride-Along	Difference
Mean Waiting Time	.0253	1.4091	1.3838
Mean Queue Length	.2094	.4948	.2854
Mean Number of Officers Used	.0062	.1381	.1319
Mean Number in System	0.7502	1.8222	1.072

Table 3: Markovian Random Server Model (8 Officers).

	Pre Ride-Along	Post Ride-Along	Difference
Mean Waiting Time	.0061	.4251	.419
Mean Queue Length	.1825	.4295	.247
Mean Number of Officers Used	.0019	.0503	.0484
Mean Number in System	.7509	1.7711	1.0202

In Table 1 in the Pre Ride-Along column, we observe that the mean wait is .054 minutes and the mean number of callers waiting is .24. In Table 2 in the Pre Ride-Along column, we observe that the mean wait is .03 minutes and the mean number of callers waiting is .21. In Table 3 in the Pre Ride-Along column, we observe that the mean wait is .01 minutes and the mean number of callers waiting is .18. From the results in these tables, we are led to believe that the IPD is doing well at handling calls as the queue length and waiting time are near zero. Moreover, it tells a picture that 7 officers might not be necessary to achieve good outcomes for the Ithaca community and 6 offices could be enough. However, this perspective is rather contradictory given the number of officer complaints about working too many hours. We will explore this further when we integrate our officer-perspective data into the simulation.

4.2 INCORPORATING RIDE-ALONG INFORMATION INTO THE SIMULATION

Using the data exclusively, we observed that our model predicted the IPD officers appeared much more idle than what they had claimed. Thus, we decided to go on several 4-hour ride-alongs with Officer Kurt Soderholm and Officer Michael Meskill. These ride-alongs were quite illuminating because we were able to experience first hand how calls were handled from start to finish. In addition, we interviewed other officers about their experiences at the IPD, allowing us to discover several aspects of policing that are not captured by the call data and are vital for successful and thorough policing. Thus, the call data was good for understanding the arrival part of the process, but not necessarily the service part of the process.

The most important observation we made from our ride-alongs was that following up on investigations/calls was a salient (and very time consuming) part of the policing process. These actions do not make it into the call data, mainly because the follow-up work that officers do is not time-stamped and recorded. Moreover, re-opening cases and compiling data on previous calls is not recorded either the call data we received only shows the initial response to calls. After interviewing several officers about possible time estimates for followup on the most frequent calls, we were able to generate Table 4, which approximates the average followup time for the top ten offenses. Unfortunately, for ASSIST, SUSPICIOUS, and TRAFFICOFFENSE, the officers we interviewed could not reach a consensus on how long they took since they were highly variable and depended heavily on the situation the officers were in. More work in this regard is needed to fully determine the appropriate distribution of these times. If the completion times were too variable we indicate in Table 4 that the follow-up time was *highly variable*.

Table 4: Table detailing the estimated times of completion for certain types of calls

Offense	**Followup Time**	**Offense**	**Followup Time**
ASSIST	25 mins (*Highly Variable*)	SUSPICIOUS	20 mins (*Highly Variable*)
NOISECMPLNT	15 mins	THEFT	2 hrs
PARKINGPROBLEM	20 mins	TRAFFICCMPLNT	25 mins
PDACCIDENT	30 mins	TRAFFICOFFENSE	30 mins (*Highly Variable*)
PROPERTYCMPLNT	2 hrs	WELFARECHECK	1 hr

As a result of our ride-alongs, we augmented our call data with the ride-along results of Table 4 to re-adjust our simulation parameter values. By adding this information, a call would not be considered complete when an officer entered a completion time, but instead a call was completed when all followup work was done. This mismatch in the two times hints at an important lesson. The data cards which we analyzed were from the perspective of the *customer*. A 911 caller has access to information such as call time, when the officer arrives, and when they leave. However, the whole story is attained when we integrate these data with *officer-perspective* data. Performance measures like the amount of followup, which are very important for analysis of the police department, but are not easily shown in the caller data. Many other qualitative things about the police department's operations **can only be learned** by interacting with officers and observing how they spend their time in a variety of situations.

4.3 SIMULATION RESULTS: OFFICER-PERSPECTIVE DATA INTEGRATED

Now, we will describe the results for the simulations with extra followup time added and officers removed. In this simulation, we add an additional 15 minute of follow-up work for each call. The results are given in the Post Ride-Along column in Tables 1 - Table 3. In Table 1 in the Post Ride-Along column, we observe that the mean wait is 3.4 minutes and the mean number of callers waiting is .57. In Table 2 in the Post Ride-Along column, we observe that the mean wait is 1.41 minutes and the mean number of callers waiting is .49. In Table 3 in the Post Ride-Along column, we observe that the mean wait is .43 minutes and the mean number of callers waiting is .43. From the results in these tables, we are led to believe that the IPD not doing as well as indicated in the Pre Ride-Along column and these results are a bit more consistent with the comments given by the officers.

We observe in Tables 1 - 3 that all of the performance measures are uniformly worse and and reflect some of the issues that the IPD has been complaining about. Moreover, we know from the queueing theory literature that removing one officer (in a small server setting) can have a large impact on performance measures. One of the most important things to consider when staffing a police department is that officers might be injured in the line of duty. Thus, it can dramatically affect the performance of the IPD and several injuries were reported at the time of this work. One complication is that injured officers are not easy to replace since officers are required to go through extensive training and preparation. Thus, it is important to staff the IPD beyond levels that meet the performance measures that would be agreed upon by IPD and the Ithaca government to make sure if an officer is injured that the IPD can still operate at a high level.

5 CONCLUSION & FUTURE DIRECTIONS

In this paper, we combine simulation and data science to construct a queueing model for modeling the dynamics of the Ithaca Police Department. The results from this research indicate that there are observable benefits to having at least six officers. It appears that this minimum of six officers tends to be adequate overall. However, we also observe that adding additional officers such as 7 or 8 can help the IPD respond even quicker to calls. Moreover, it is clear that the additional officers can alleviate stress, injury, and overtime considerations that are hard to mathematically quantify. Removing or adding one officer can have a big impact on the IPD operations and the success of the IPD.

There were several limitations to this analysis that are worth considering. First, there was no hard data available to us on followup work, which turns out to be an important factor in assessing staffing levels. Second, it is difficult to say how many officers should be on payroll with the information we have been given. Specifically, we are unable to comment on whether or not this shift minimum should be reached by utilizing overtime with current personnel or by hiring additional officers to fill shifts permanently. In deciding which of these methods to pursue, one must weigh a great amount of data that was simply not available to us, such as the financial implications of overtime rates vs. benefits for new hires, any causal relationships between different amounts of overtime and performance decline or staff retention, and availability of candidates in the appropriate areas. Hopefully, though, the data presented here will illuminate this side of the decision making.

To further extend this project, it would be beneficial to look at more nuanced elements of police operations. Possible avenues include examining how staffing changes can be made according to the time of day as well as the time of year. It would then be possible to weigh the performance of this dynamic staffing against the direct costs of a constant staffing level. It would also be beneficial to consider how best to incorporate overtime hours worked into our assessment of a staffing level's effectiveness.

In a broader sense, our data-driven simulation framework is especially useful given the current climate, where funding levels for policing is the subject of intense public discussion. In fact, the Mayor Svante Myrick of Ithaca recently announced plans to replace the IPD with a civilian-led agency. He proposed a new entity called the Department of Community Solutions and Public Safety. This agency would include both armed and unarmed first responders that are trained to de-escalate a number of situations, all of whom will report to a civilian public safety director. Our simulation and analysis is essential to understanding the efficacy of this new entity. Our simulation is robust in that it allows officials to assess the police department's performance for a wide variety of staffing levels and could also be used for evaluating the performance of this new agency with appropriate staffing levels. It is our hope that our work will be used to help guide Ithaca's officials in making decisions about future resource allocation for IPD or any new public safety entity.

ACKNOWLEDGEMENTS

We would like to thank Engineering Learning Initiatives at Cornell for providing Matthew Ziron with support for conducting this research. Christopher Archer would like to thank the Rawlings Presidential

Research Scholars program for their support to do this research. Finally, we would like to give our sincere thanks to the Ithaca Police Department, especially Sergeant Tomberelli, for providing us with the data and their time to conduct the research.

REFERENCES

"A Queueing System in which Customers Require a Random Number of Servers, author=Green, Linda, journal=Operations research, volume=28, number=6, pages=1335–1346, year=1980, publisher=INFORMS".

Brill, P. H., and L. Green. 1984. "Queues in which Customers Receive Simultaneous Service from a Random Number of Servers: A System Point Approach". *Management Science* 30(1):51–68.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. 2005, March. "Statistical Analysis of a Telephone Call Center". *Journal of the American Statistical Association* 100(469):36–50.

Daw, A., R. C. Hampshire, and J. Pender. 2019. "Beyond Safety Drivers: Staffing a Teleoperations System for Autonomous Vehicles". *arXiv preprint arXiv:1907.12650*.

Daw, A., and J. Pender. 2018a. "Exact Simulation of the Queue-Hawkes Process". In *Proceedings of the 2018 Winter Simulation Conference*, 4234–4235.

Daw, A., and J. Pender. 2018b. "Queues Driven by Hawkes Processes". *Stochastic Systems* 8(3):192–229.

Federgruen, A., and L. Green. 1984. "An M/G/c Queue in which the Number of Servers Required is Random". *Journal of applied probability*:583–601.

Green, L. 1981. "Comparing Operating Characteristics of Queues in which Customers Require a Random Number of Servers". *Management Science* 27(1):65–74.

Green, L. 1984. "A Multiple Dispatch Queueing Model of Police Patrol Operations". *Management Science* 30(6):653–664.

Hampshire, R. C., S. Bao, W. S. Lasecki, A. Daw, and J. Pender. 2020. "Beyond Safety Drivers: Applying Air Traffic Control Principles to Support the Deployment of Driverless Vehicles". *PLoS one* 15(5):e0232837.

Ignall, E. J., P. Kolesar, and W. E. Walker. 1974. "The Use of Simulation in the Development and Empirical Validation of Analytic Models for Emergency Services". In *Proceedings of the 7th conference on Winter simulation-Volume 2*, 529–537.

Kahara, M., and J. Pender. 2017. "Using Simulation to Study the Impact of Racial Demographics on Blood Transfusion Allocation Policies". In *2017 Winter Simulation Conference (WSC)*, 2786–2797. IEEE.

Massey, W. A., and J. Pender. 2018. "Dynamic Rate Erlang-A Queues". *Queueing Systems* 89(1-2):127–164.

Nirenberg, S., A. Daw, and J. Pender. 2018. "The Impact of Queue Length Rounding and Delayed App Information on Disney World Queues". In *2018 Winter Simulation Conference (WSC)*, 3849–3860. IEEE.

Novitzky, S., J. Pender, R. H. Rand, and E. Wesson. 2019. "Nonlinear Dynamics in Queueing Theory: Determining the Size of Oscillations in Queues with Delay". *SIAM Journal on Applied Dynamical Systems* 18(1):279–311.

Palomo, S., J. Pender, W. Massey, and R. C. Hampshire. 2020. "Flattening the Curve: Insights From Queueing Theory". *arXiv preprint arXiv:2004.09645*.

Pender, J., S. Tao, and A. Wikum. 2020. "A Stochastic Model for Electric Scooter Systems". *Available at SSRN 3582320*.

Schaack, C., and R. C. Larson. 1989. "An N Server Cutoff Priority Queue where Arriving Customers Request a Random Number of Servers". *Management Science* 35(5):614–634.

Verma, A., R. Ramyaa, S. Marru, Y. Fan, and R. Singh. 2010. "Rationalizing Police Patrol Beats using Voronoi Tessellations". In *2010 IEEE International Conference on Intelligence and Security Informatics*, 165–167. IEEE.

Zeltyn, S., and A. Mandelbaum. 2005. "Call Centers with Impatient Customers: Many-Server asymptotics of the M/M/n+ G Queue". *Queueing Systems* 51(3-4):361–402.

AUTHOR BIOGRAPHIES

CHRISTOPHER ARCHER is an undergraduate student in Operations Research and Information Engineering at Cornell University. His research interests are in queueing theory, machine learning and stochastic simulation. His e-mail address is caa234@cornell.edu.

MATTHEW ZIRON is an undergraduate student in Operations Research and Information Engineering at Cornell University. His research interests are in queueing theory, machine learning and stochastic simulation. His e-mail address is maz63@cornell.edu.

JAMOL PENDER is an associate professor in the School of Operations Research and Information Engineering (ORIE) at Cornell University. He earned his PhD in the Department of Operations Research and Financial Engineering (ORFE) at Princeton University. His research interests include queueing theory, stochastic simulation, dynamical systems and applied probability. His e-mail address is jjp274@cornell.edu. His website is https://blogs.cornell.edu/jamolpender/.