Modelling the Spread of Complaints Among Police Officers

Jeremy Hudson Chan

Honors thesis presented for the degree of
Bachelor of Arts (Honors) in Computer Science

Morrissey College of Arts & Sciences
Department of Computer Science
Boston College
Chestnut Hill, MA
United States
May 6, 2023

Advisor: Prof. George Mohler Ph.D.

Abstract

Recent research has proposed that police behavior, including excessive use of force and misconduct, can spread through a contagion process on police social interaction networks. However, the literature is mixed on this topic, where null findings have followed positive findings of contagion in some cities in others. Robust and easy-to-implement statistical tests are needed to help police agencies detect the contagion of unwanted police behavior in policing networks. In this thesis, we contribute to this literature on whether excessive police use of force is socially contagious in officer networks and present several models of excessive use of force complaints. Our proposed algorithms do not require as many parametric modeling assumptions as existing models. We propose a set-based Resampling Test variation as a baseline model to test the hypothesis that the use of force arises through a random, independent event process. We then compare this baseline to two other algorithms we adapt for modeling police use of force: the Pólya Urn Model and Spillover Pólya Urn Model. Finally, we implement these three algorithms using police use of force data from Chicago. Our results indicate that contagion plays a role in the excessive use of force in Chicago police networks.

Acknowledgements

Words cannot express my gratitude to my advisor Prof. George Mohler for guiding me through writing this Honors Thesis. Without you, I would have been stuck at step one, trying to find a topic. It was a pleasure to work for you and learn about the science behind social contagion, criminology, and research.

Special thanks to my Dad and Mom for motivating me when times were tough and giving me pointers on what to improve. Hopefully, I have made you proud.

I'd like to acknowledge Yuting Ji for suggesting random ideas on what I could add to this thesis and making learning way more interesting and fun. As for my other friends and family, thank you for motivating me to finish; I would not have done it without you (you know who you are).

List of Figures

List of Tables

Contents

Chapter 1

Introduction

In recent years, there have been more and more publicized incidents of excessive use of force by the police; an example of this would be the killing of George Floyd on May 25, 2020. [1] More incidents of excessive use of force and police brutality can be found in The New York Times and are reported all over news stations and social media. [2] Following such highly publicized incidents, it is unfortunate that the relationship between the police and the public seems to have become increasingly tenuous [3]. Last year, roughly 600,000 to 700,000 U.S. residents aged 16 and older claimed to have experienced police contact involving less-than-lethal force. In addition, approximately 75% to 83% of U.S. residents reported that police used excessive force [3].

There has been research that suggests that the spread of violence and diseases are parallel to each other. Some papers have investigated how violence is similar to infectious diseases and present better guidance for future strategies for reducing violence [4]. Based on this theory, much research is done on top of investigating the social contagion of gunshot incidents, gang activities, the adolescent age group, and many more [5, 6, 7]. We will look into the specifics and findings of these results in Chapter 2 Literature Review.

Moreover, in recent years, the investigation of social contagion in police networks and its role in the spread of police misconduct has risen. There is an ongoing debate on whether social contagion plays a role in these excessive use-of-force cases. [8, 9, 10] Both sides are at odds on whether social contagion plays a role. To contribute to the discussion, we must first consider how these excessive use-of-force cases occur. Are these merely just randomized events happening to random officers and citizens?

In this thesis, we derive algorithms aiming to simulate and model the distribution of use-of-force complaints of police officers. We will consider the factor of social contagion in our models to better determine whether it impacts the closeness of our simulated distributions. We believe that these models could serve as a benchmark of whether the current police misconduct justice system serves its purpose in reducing violent police incidents and allegations against them. Our models would also contribute to the current argument of whether police excessive use of force is socially contagious.

Therefore, given the prior research on violence potentially being a social contagion, we will attempt to incorporate this idea into our algorithms and mathematically represent this in our study. To set the premises, we define social contagion here as the spread of effect, attitude, or behavior from an initiator to a recipient where the recipient does not perceive an intentional influence attempt on the part of the initiator [11].

We propose three algorithms, Resampling Test, Pólya Urn Model, and Spillover Pólya Urn Model. Our results show that the Spillover Pólya Urn Model performs best, followed by the Pólya Urn Model and then the Resampling Test. We believe that the Pólya Urn Model can serve as a baseline for the Police Force in the future to judge whether the general force has fewer officers who receive multiple complaints in their tenure. The Spillover Pólya Urn Model can serve as a good simulation to simulate current Police Use-of-Force complaint numbers in other areas.

Our simulations of the proposed algorithms were run based on the Invisible Institute: Citizens Police Data Project (CPDP) dataset. The dataset keeps track of the allegations that the Greater Chicago Police Force receives from 2000 to 2016. We assume that all the allegations received by the Police Force and Invisible Institute are accurate and not false. We chose this dataset due to its visibility and easy-to-understand interface, as well as its being in a larger city, providing us with potentially more data.

Our proposed methods can benefit police departments by helping them predict the number of complaints they might receive. In addition, they could use our algorithms as a risk assessment to see whether police violence has decreased. More of our ideas will be presented in Chapter 6: Conclusion & Future Work.

Chapter 2

Literature Review

2.1 Reasonings behind Police Violence

Before we can develop mathematical models to simulate allegation numbers, we should consider the perspective of police officers and the reasoning behind their actions, whether they are using excessive force.

The National Institute of Justice published a study in May 2000 showing that in 1996, improper use of force was used in 38 percent of encounters involving force. [12] Back in 1967, 35 percent of meetings involving force were excessive usage. The data here were drawn from medium-sized sample sizes of 5688 and 1565 cases, respectively. Many believed that though these incidents were not common, they should be permitted to use more force than the law back then, as they thought that, in some cases, it was acceptable for more force to be applied. [12] Other literature mention how higher education in officers leads to less abuse of authority from police officers. [13] Aside from these papers, little research has been conducted on police officers' stances before committing excessive force.

There has been a lot of research on police brutality and its intersection with racial bias. Sharara and colleagues published a paper in October 2021, with their findings showing that out of 30800 deaths from police violence between 1980 and 2018, the age-standardized mortality rate due to police violence was highest in non-Hispanic Black people (0.69 per 100000), with other races having a significantly lower mortality rate. [14] The age-standardized mortality rate has significant discrepancies between states, increasing by 38.4% throughout the study. Many other studies have researched the intersection between racism and police brutality and have either concluded a specific connection between race and police brutality or advocated for more research on the subject. [15, 16, 17]

These studies highlight the importance of allocating more resources to prevent these incidents and that more research should be done. With the mortality rate increasing throughout the years and the highly publicized incidents of police brutality intersecting with racial bias, more preventive methods must be considered. Lersch and Mieczlowski, in their paper, suggested future research directions such as the need for more national studies, clear definitions of violent behavior by officers, and more statistical analysis. [18] More psychological, sociological, and organizational theories are proposed in the paper, as well as solutions to reduce the rate of violent police misconduct. As this paper was published in 2004, we will look into more literature related to police violence.

2.2 Predicting Police Violence

Many studies have investigated the factors and conditions that influence these numbers. These factors include the presence of a full-time internal affairs department, in-service training, violent crime rates in the area, academy performance, off-duty incidents, etc. [19, 20] The problem with these research papers is that the data variables that they suggest to whether an officer would commit are either not freely available to the public, or are they even tracked at all. Therefore, while the variables indicated in their papers are essential in theory, we cannot verify whether they are consistent over different locations, times, and datasets.

There has been some research on predicting the number of complaints using the origin of complaints. Greene and colleagues used the statistical method Odds Ratio (OR), a measure of association between an exposure and an outcome, to identify correlated variables.[20] This method helps identify variables, but when in the case where these variables are not freely available, it would not be able to estimate the total number of cases. Furthermore, predicting whether an individual would use excessive force is also rather intrusive; hence we decided to pursue other methods which involve less specific variables.

Some research has also been done on whether complaints against officers are successful and sustained. [21] Griswold uses multivariate analysis to conclude that more severe complaints, such as excessive force and battery complaints, are more likely to be sustained. The literature here helps us comfortably assume that we can look at the number of excessive use of force complaints to predict actual violence cases. These complaints seem to imply force deemed unreasonable by the police force.

These papers mentioned in this section mainly use different statistical methods; recently, more research papers have investigated the social networking and social contagion side of violence, which we will review in the next section.

2.3 Social Contagion as a Factor in Violence

In Slutkin's Book Cure Violence: Treating Violence As a Contagious Disease, he explains the theory behind seeing violence as an infectious disease. Slutkin describes violence as an infectious population disease, including clustering, spread, and transmission factors. He also explains how violence also shows the characteristics of an infectious disease in an individual, including factors such as susceptibility, exposure, incubation, and many more. Slutkin then concludes that treating violence as an infectious epidemic would be an effective solution by (1) detecting and interrupting ongoing and potentially new infectious events; (2) determining who is most likely to cause further infectious events from the infected population and then reducing their likelihood of developing disease and subsequently transmitting; and (3) changing the underlying social and behavioral norms, or environmental conditions, that directly relate to the spread of the infection. [4] In our experiments, we took inspiration from strategies (1) and (2), as we believe that these two strategies could lead to more of an immediate impact compared to strategy (3), as well as the limitations of our data and resources.

Papachristos has done a lot of research into violence as a contagious disease. In Papachristos' paper "The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence," he investigates how geography and social networks affect gang violence in Chicago and Boston. Results show that adjacency of gang turf and prior conflict between gangs are strong predictors of subsequent gang violence, with social factors such as reciprocity and status-seeking also contributing to gang violence. [5] There have been many news reports about police gangs in Los Angeles. Could the findings in Papachristos' paper here be somewhat related to that? Papachristos has also written other papers that relate to social contagion and its relationship to gunshot violence. [6, 7]. These papers also clearly show the significance of one's social network when encountering gunshot violence. We must consider whether an officer's social network could also influence excessive use of force numbers.

2.4 The Debate on Whether Social Contagion is a Factor in Police Networks

Papachristos published papers in 2019 studying the effect of an officer's exposure to peers previously accused of misconduct and recreated the network structure of the misconduct contagion. [8, 9] The methods Papachristos used involved creating a misconduct network and conducting Bayesian statistics on factors affecting the number of complaints received. His results show that officer involvement in excessive use of force complaints is predicted by having a more significant proportion of co-accused with a history of such behaviors. [8] While constructing the network, he found that police misconduct is associated with attributes including race, age, and tenure and that almost half of the police officers are connected in misconduct ties in broader networks of misconduct. [9] They also found that certain dyadic factors, especially seniority, and race, strongly predict network ties and the incidence of group misconduct. [9] Papachristos supported his argument by applying the growing field of network science to explore whether police violence would be associated with characteristics of an officer's social networks and placement within those networks. [22] His findings suggest that policies and interventions aimed at curbing police shootings should include individual assessments of risk and an understanding of officers' positions within larger social networks. [22]

Chalfin, in 2021, countered Papachristos' findings by removing the "bad apple" officers from police networks led to an insignificant reduction of police misconduct (4% - 6%) in only the top 10% of officers based on ex-ante if middle-risk officers replaced them. They suggest that surgically removing predictably problematic police officers is unlikely to impact citizen complaints significantly. Chalfin provides support for the idea that early warning systems must be designed, above all, to deter problematic behavior and promote accountability. [10]

Papachristos responded to Chalfin's study by releasing his interpretations of Chalfin's results. He believes that Chalfin's results are squarely within the range of other interventions to reduce police complaints and use of force. Once network spillovers are accounted for, estimates are up to five times as large. [23] Papachristos argues that removing problem officers is a normative good that should be pursued on moral grounds. [23]

In 2022, Simpson and Kirk published findings based on Dallas, Texas data, indicating that the risk of a Dallas PD officer engaging in misconduct is not associated with the disciplined misbehavior of her ad hoc, on-the-scene partners. [24] Rather, a greater risk of misconduct is associated with past misbehavior, officer-specific proneness, the neighborhood context of patrol, and, in some cases, officer race, with departmental tenure as a mitigating factor. [24] They suggest that actor-based and ecological explanations of police deviance should not be summarily dismissed in favor of accounts emphasizing negative socialization. They raise the possibility that results are partly driven by unobserved trait-based variation in the situations that officers find themselves in. All in all, interventions focused on individual officers, including the termination of deviant police, may be fruitful for curtailing police misconduct—where early interventions focused on new offenders may be key to avoiding the escalation of deviance. [24]

From the above literature, we can see whether social contagion plays a role in police excessive use of force, is being debated in recent years. There has not been any directly related research modeling the distribution of complaints among police officers. Most research has been done on the different factors of police misconduct, gang violence, and the effect of social networks. Not many provide benchmark methods to predict the number of use-of-force complaints that a group of officers might receive. Our experiment supports the idea of social contagion being prevalent in police networks and brings awareness to the social contagion aspect of behaviors around us.

Chapter 3

Methodology

3.1 Resampling Test

3.1.1 What is the Resampling Test

Firstly, let us create a null hypothesis on effectively modeling the distribution. The null hypothesis is that some officers, by random independent pairings, will have differing amounts of use of force incidents. To explain, let us ignore all other factors and approach this problem from a mathematical angle. If an incident is at a particular location (near a beat), they randomly send out several officers to respond. Out of these police force responses, the ones receiving complaints will be noted and added to the data. How would we simulate this?

Given that an incident happens near a police beat, let us set the number of officers per incident to the original number of officers in each incident, as there is no method for us to retrieve and simulate the number of officers per incident. Now, we can simulate which officers receive complaints from each incident by uniformly selecting them from each Beat involved. We can simulate this by making all the officers' weights proportional to the number of officers in the beats.

The idea here is that, for each incident, we get the original Beat where the incident occurs. We would then do sampling without replacing the number of officers involved in the incident (k). When an officer is sampled in an incident, we will increment their count by 1 to keep track of how many times they are selected throughout the process. To summarize, the main idea of the Resampling test is to see if officers with a high number of incidents could occur by random choice. We will reject the null hypothesis if the produced distribution differs from our original data distribution.

The following subsection should give a more step-by-step method to understand the idea better, while the last subsection will be the pseudocode implementation of the Resampling Test.

3.1.2 The Resampling Test as an Algorithm

Step 1: Initialize each officer's chance at being selected for an incident by proportioning them over the Total Number of Officers from each Beat (T) (if Beat 1 has

100 officers, the starting weight would be 0.01):

$OfficerWeight = \frac{1}{T}$

Step 2: Select an unselected incident, get the number of officers (x) in the original incident, and get the Beat of where they are from.

Step 3: Select officers from the same Beat of the original incident, with each officer's probability of being selected adjusted according to their weight.

Step 4: Add 1 to each officer's total count of complaints, each being part of the selected x officers.

Step 5: Repeat Steps 2 to 4 until all incidents are accounted for.

3.1.3 Pseudocode

Algorithm 1 Resampling Test (Uniform Unchanged Weights)

```
for beat in beats do
    for officer in beat do
        {Initialize weights of officers in each Beat to be uniform}
        officer[weight] = 1 / len(beat)
    end for
end for
for incident in incidents do
    k = len(incident) {k = number of officers in each incident}
    beat = incident.beat {get the Beat of the officers involved in the incident}
    beatOfficers = beat.officers {list of all officers in the incident beat}

    randomOfficers = random.choose(beatOfficers, k, weight)
    for officer in randomOfficers do
        officer[count] += 1 {update the number of complaints of an officer by 1}
    end for
end for
```

3.2 Pólya Urn Model

3.2.1 What is the Pólya Urn Model

We must realize that in real life, not everyone is equally likely to receive complaints while on duty. Also, as mentioned before, research shows that many police officers believed excessive force could sometimes be acceptable. [12] The previous model assumes that the probability of an officer receiving a complaint is constant throughout, but this should not be the case. So, how should we incorporate the different probabilities of each officer? Officers who believe excessive force is acceptable will tend to replicate the behavior at a higher probability, while officers who do not believe in it will have a lower probability. Here, we propose the Pólya Urn Model/Chinese Restaurant Process algorithm to simulate Police Use-of-Force complaints.

Here, we introduce a new parameter, alpha (α), the constant we will add to each officer after receiving a Use-of-Force complaint. As mentioned in Chapter 2 Literature Review, some officers believe that excessive use of force is necessary when they deem so, even though it might be against the law. We implement this idea by adding a constant (α) to the weight proportion of an officer. If we increment each officer by α each time they receive a complaint, they will likely be selected in future complaint incidents in our simulation.

3.2.2 The Pólya Urn Model as an Algorithm

Step 1: Initialize each officer's chance at being selected for an incident by proportioning them over the Total Number of Officers from each Beat (similar to 3.1)

Step 2: Select an unselected incident, get the number of officers (x) in the original incident, and get the beat of where they are from.

Step 3: Select officers from the same beat of the original incident, with each officer's probability of being selected adjusted according to their current weight.

Step 4: Add 1 to each officer's total count of complaints, with each officer being part of the selected x.

Step 5: Add α to each officer's weight that was part of the selected x (same officers in Step 4).

Step 6: Repeat Steps 2 to 5 until all incidents are accounted for, with each iteration using the newly updated officer weights.

3.2.3 Pseudocode

Algorithm 2 Pólya Urn Model (Constantly Changing Weights)

```
for officer in beat do
    officer[weight] = 1 / len(beat)
end for
for incident in incidents do
    k = len(incident)
    beat = incident.beat
    beatOfficers = beat.officers
    randomOfficers = random.choose(beatOfficers, k, weight)
    for officer in randomOfficers do
        officer[weight] += α{weights will now be updated to new values}
        officer[count] += 1
    end for
end for
```

3.3 Spillover Pólya Urn Model

3.3.1 What is the Spillover Pólya Urn Model

After the Pólya Urn Model above, we can now consider additional factors that affect the distribution of Police Use-of-Force complaints. One factor that has been investigated thoroughly and was mentioned in Chapter 2 Literature Review is violence as a disease. In the case of gang violence, if gang members grow up in an environment where violence is the norm, with violent behavior being the answer to all problems, they will be more inclined to react similarly. To theorize, if officers see fellow officers in Use-of-Force incidents, would they be affected and accustomed to the idea? Therefore, we must consider how the idea of behavior spreading between humans affects the distribution of Police Officers' Use-of-Force complaints.

How should we incorporate this into our current working Pólya Urn Model? Here, we propose an equation with an additional parameter, Beta (β). β is the proportion of the weight change that a Police Officer could receive a complaint again in the distribution, with $1 - \beta$ being the proportion of an officer's weight changes due to other officers involved in a complaint. To better explain this, we have proposed a formula for updating officer weights when involved in an incident. The formula is shown below:

Formula 3.3.1: $updateOfficerWeight = \beta \cdot (w_i + \alpha) + \frac{1-\beta}{x-1} \cdot \sum_{j \neq i}(w_j + \alpha)$

Formula 3.3.1 encapsulates the idea of social contagion mathematically. First, we update each officer's weight by α, as we did previously. Then, we multiply that with β, which is the proportion that the officer's new weight will. Then, the second part of the formula takes the average of all officers' weight that the officer had worked with in that allegation event and multiplies it with $1 - \beta$. The new officer's weight will now either be increased or decreased as their weight could be the highest out of everyone, leading to their weight being lowered due to the other officers having lesser weights (fewer historical complaints).

3.3.2 The Spillover Pólya Urn Model as an Algorithm

Step 1: Initialize each officer's chance at being selected for an incident by proportioning them over the Total Number of Officers from each Beat (T) (similar to 3.1 & 3.2)

Step 2: Select an unselected incident, get the number of officers (x) in the original incident, and get the beat of where they are from.

Step 3: Select officers from the same beat of the original incident, with each officer's probability of being selected adjusted according to their current weight.

Step 4: Add 1 to each officer's total count of complaints, with each officer being part of the selected x.

Step 5: Add α to each officer's weight that was part of the selected x (same officers in Step 4).

Step 6: Multiply the updated weight (oldWeight + α) by β, then multiply $1 - \beta$ with the sum of all the updated weights of the other officers divided by x – 1 (**Formula 3.3.1**).

Step 7: Repeat Steps 2 to 6 until all incidents are accounted for, with each iteration using the newly updated officer weights.

3.3.3 Pseudocode

Algorithm 3 Spillover Pólya Urn Model (Constantly Changing Weights)

```
for beat in beats do
    for officer in beat do
        officer[weight] = 1 / len(beat)
    end for
end for
for incident in incidents do
    k = len(incident)
    beat = incident.beat
    beatOfficers = beat.officers
    randomOfficers = random.choose(beatOfficers, k, weight)
    for officer in randomOfficers do
        otherOfficerWeights = sum(weight of officers != officer)
        officer[weight] = β · (officer[weight]+α) + (1−β)/(k−1) · (otherOfficerWeights)
        officer[count] += 1
    end for
end for
```

Chapter 4

Data

4.1 Background & Data

4.1.1 Invisible Institute: Citizens Police Data Project (CPDP)

The Invisible Institute is a nonprofit journalism production company based in Chicago, Illinois, that investigates and holds public institutions accountable. They employ numerous tactics such as investigative reporting, multimedia storytelling, human rights documentation, etc. Their work centers around a central principle: "We as citizens have co-responsibility with the government for maintaining respect for human rights and, when abuses occur, for demanding redress." [25]

More information about the Invisible Institute and its history, operations, and achievements is on its website https://invisible.institute/.

The Invisible Institute has multiple open-source datasets widely available to the public. The dataset we used to investigate here is the Citizens Police Data Project (CPDP). A press release regarding the dataset explaining initial observations and a brief description of the dataset can be found at the following link: https://invisible.institute/press-release. The actual dataset itself and be found at the following link: https://cpdp.co/

To summarize, the CPDP dataset records police interactions with the public. It opens them up to make the data applicable to the public, thus creating a permanent record for every Chicago Police Department (CPD) police officer who has a complaint against them. More specific information and data usage will be given in Chapter 4.2, Dataset.

4.1.2 A Description of the Data

The dataset is visualized on the website mentioned above (https://cpdp.co/) by an interactive geographical map. You can individually search queries which include geographical, complaint categorical, outcomes of complaint investigations, complaint characteristic, complaint demographic, officer demographic, and officer characteristics.

In our scope of the investigation, we decided to take more of a temporal-spatial approach; hence the data we collected initially to set up our distribution of officers was dependent on time. As a result, the data we collected was in the time frame between 2000 to 2016, which is the time frame of what other literature has followed. [26]

The data format we collected from the website was Excel; we had one file each year. Each Excel file consists of 4 spreadsheets; Allegations, Police Witnesses, Complaining Witnesses, and Officer Profile. The Allegations spreadsheet details the information of each allegation that an officer received that year. Police and Complaining Witnesses contain information about witnesses behind each allegation case (not all cases have witnesses). Finally, the Officer Profile spreadsheet includes the information of each officer by their Officer ID. In addition, their Names, Gender, Race, Appointment Dates (Excel Numeric Date), Ranks, and Ages are included.

For our investigation, we mainly focused on the first and last type of spreadsheets in each Excel File (Allegations and Officer Profile), aiming to simulate the distribution of allegations received by officers. The data preprocessing steps will be given in detail in the following chapter, Chapter 4.2 Data Preprocessing.

4.2 Data Preprocessing

4.2.1 Assumptions

As mentioned on the CPDP website's disclaimer, the accuracy of the information provided by CPDP cannot be guaranteed. CPDP instead commits itself to being honest about flaws and transparent in its publishing process. In addition, the code they used in their preprocessing is publicly available in the following GitHub Repository: https://github.com/invinst/chicago-police-data. Hence, our assumptions about the data align with CPDP's disclaimer.

4.2.2 Steps

As mentioned above, the main spreadsheets needed for our experiment are the Allegations and Officer Profile spreadsheets. We first proceeded to combine all of the allegations from 2000 to 2016 into one file as it would ease the process of tallying the number of allegations per officer. Next, we filtered out allegations unrelated to police violence or excessive use of force and disposed of that data. Then we took the tally of the number of allegations per officer by counting the number of occurrences of each Officer ID.

Afterward, we combined the Officer Profile spreadsheets into one file while removing the duplicate officers. Doing this would give us an accurate appendix of officers with at least one allegation in the Greater Chicago Area during 2000 - 2016. We then mapped each officer in the extensive allegations file to their respective beat, which is critical in conducting our proposed algorithms above. Finally, we completed our data preprocessing with a file that contains Officer IDs, their individual Beat number, and their number of allegations.

To better represent the data after our preprocessing steps, we have summarized it into a few tables and a figure below. As a side note, there were a few instances where

1 or 2 officers had 15+ allegations against them; as these instances our outliers, we decided to group them into a category of 10+ to ease the visualization and simulation process.

Num Allegations	Officers
1	3593
2	1734
3	1006
4	658
5	404
6	264
7	193
8	119
9	111
10+	246

Table 4.1: Initial Distribution of Values

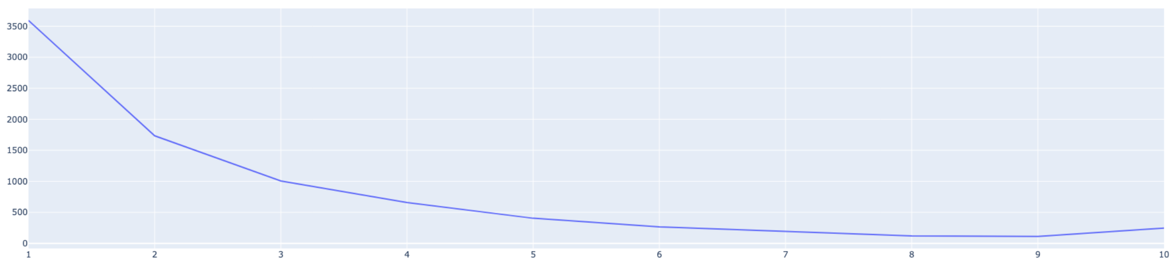

Figure 4.1: Original Distribution of Officer Complaints

Chapter 5

Experiment & Results

5.1 Experiment

For each of our proposed algorithms, we 1000 simulations for each of our algorithms and averaged the results. We used the Python programming language, the Pandas, NumPy, and Plotly libraries to implement our proposed algorithms.

For the algorithms Pólya Urn Model and Spillover Pólya Urn Model, we had the parameters α and β, which we had to find an optimal value for. First, using the Mean Square Error loss function, we took values between 0.00 to 1.00 for α and β and conducted 1000 simulations per value. Then, we took the values that would produce the lowest Mean Square Error Value, and our experiment concluded that α and β values were 0.20 and 0.40, respectively. The results and conclusions we make regarding our findings will be used under these selected values.

We included a 95% confidence interval for the mean of each complaint interval, based on the simulation results. The confidence intervals represent how ”good” our simulation estimates are and show us how our experiment is limited.

Our code implementation, data preprocessing, and results can be found at the following GitHub Repository:

https://github.com/Jeremyhudsonchan/Undergraduate-Thesis

5.2 Results

5.2.1 Resampling Test Results

Num Allegations	Officers
1	569
2	1502
3	1998
4	1798
5	1232
6	690
7	326
8	136
9	51
10+	25

Table 5.1: Resampling Test Simulation Distribution of Values

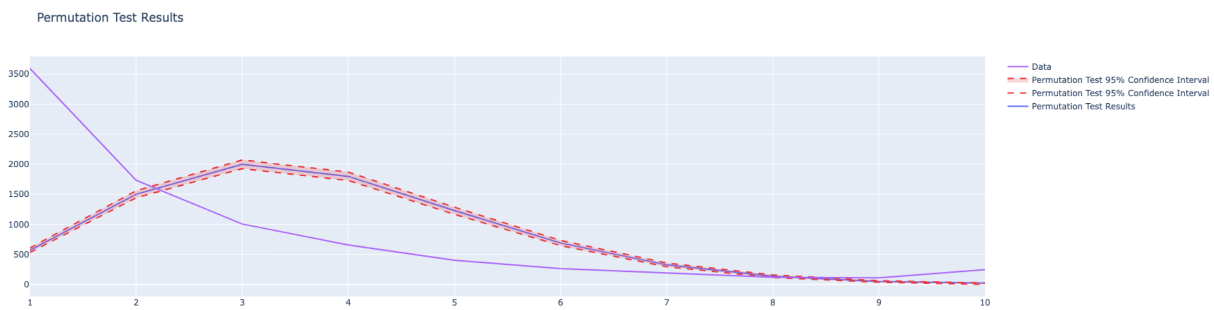

Figure 5.1: Resampling Test Results

Our results (Table 5.1 & Figure 5.1) show that the Resampling Test algorithm does not encapsulate the distribution trend of police officer allegations. As you can see, if we were to follow the Resampling Test Results, we would expect Officers who have one allegation filed against them, most of them we go on to receive two more allegations against them (hence the maxima at the third interval). But when we compare this to the original data distribution, the maximum of the data remains at one allegation (when an officer receives a complaint, they will not receive any more complaints during their tenure).

The findings support the argument of rejecting the initial null hypothesis, suggesting that independent random pairings of officers are insufficient to explain the distribution. The method suggests that too many officers repeat their violent actions and over-estimates the number of officers with three to seven allegations. Therefore, we need to consider different factors that an officer would repeat excessive use of force (i.e., their beliefs in when to use what force).

In reality, we believe it would not be suitable to use this to model the distribution of police officers. For the reasons mentioned above, it would grossly overestimate the number of police officers who would be repeat offenders. This leads us to the results of our other proposed algorithms.

5.2.2 Pólya Urn Model Results

Num Allegations	Officers
1	4925
2	767
3	465
4	340
5	266
6	216
7	177
8	128
9	51
10+	20

Table 5.2: Pólya Urn Model Simulation Distribution of Values

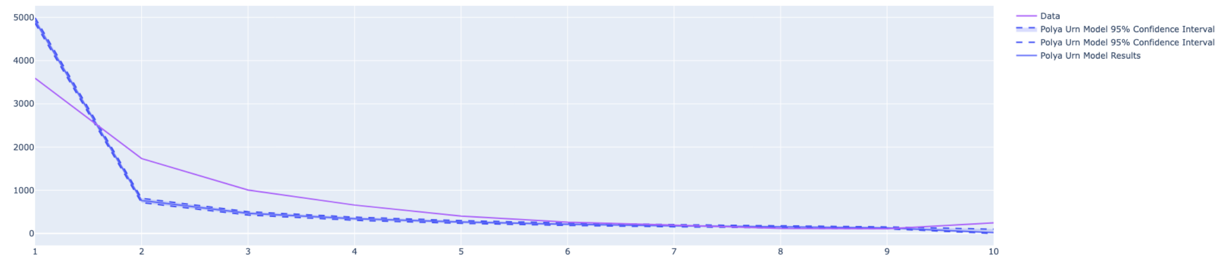

Figure 5.2: Pólya Urn Model Results

Our results (Table 5.2 & Figure 5.2) show that the Pólya Urn Model algorithm performs significantly better than the Resampling Test and better encapsulates the distribution of police officer allegations from the original dataset. Here, the total number of officers who do not have repeat allegations against them after their first complaint is around 5000. The number of officers dropping significantly for more than one allegation, only rising to 10+ allegations. Comparing this to the original distribution, we can see that this trend is similar, but the drop is insignificant. The Pólya Urn Model mainly underestimates the number of allegations, and the confidence interval range does not include the original data range.

The findings here have suggested that the model fits better than the Resampling Test, which aligns with our initial expectations that it would perform better. In reality, most officers are not inherently violent and will likely not receive multiple complaints in their tenure. Depending on the area, different police stations have other methods of handling complaints; officers might receive warnings when complaints are successfully filed against them. This also contributes to reducing officers with more than one allegation against them. We believe the model fits better than randomly sampling officers without replacement per complaint.

We propose this solution as a baseline to judge whether officers get fewer complaints over time. The distribution could be set as a target that all police departments should aim for collectively.

5.2.3 Spillover Pólya Urn Model Results

Num Allegations	Officers
1	3330
2	1151
3	761
4	608
5	505
6	417
7	342
8	277
9	222
10+	39

Table 5.3: Spillover Pólya Urn Model Simulation Distribution of Values

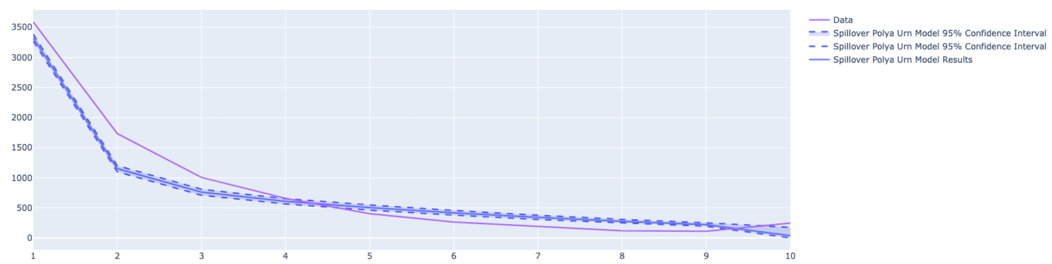

Figure 5.3: Spillover Pólya Urn Model Results

From our results (Table 5.3 & Figure 5.3), we can see that after adding our proposed formula to our code, while setting the β parameter to be 0.40, our simulation distribution models the actual distribution of values relatively closer to our other two solutions. The trend of officers from the simulation follows the actual values, which follows our hypothesis of including the social contagion factor to be the best at modeling the distribution.

As explained in Chapter 2: Literature Review, there has not been much research done on predicting officers' excessive use of force incident numbers; most previous research has only been done based on different variables which not all police departments keep track of. Based on our limited data and variables, we believe that our models simulate the number of complaints police departments receive in the Greater Chicago area over 16 years (2000-2016).

We propose this solution as an initial guide on how many excessive force complaints an area should receive over a certain period. This forecasts the potential number of future complaints, and police departments can see whether their current policies perform better or worse than this simulation, allowing them to evaluate the effectiveness of their existing system. To see a combined figure of all three simulations, as well as the Mean Squared Error for each, please refer to Appendix A: Combined Results

Chapter 6

Conclusion & Future Work

6.1 Conclusion

Our experiment's results show that the Spillover Pólya Urn Model, which mathematically incorporates the social contagion aspect of the spread of excessive force between police officers, has the best results. We believe this is one step to proving that violent acts can also be spread between officers. This is no different from other types of violence, such as gang and gunshot violence. Though our experiment's results aligned with our hypothesis, our investigation also had limitations.

Firstly, we only used one dataset of complaints in our experiments, as other areas do not provide transparent data like the Invisible Institute's Citizens Police Data Project. Secondly, our investigation was only conducted in a one-time frame from 2000 to 2016. We should have more data (say, until 2022) to see if our proposed model fits well. Lastly, our model only simulates the number of complaints received by officers given the size of the complaints (i.e., three officers received complaints in one incident). We will need to come up with a model to be able to simulate incident and complaint sizes as well. Suggestions to these limitations will be explained more in detail in Chapter 6: Future Work.

6.2 Future Work

We believe that more attention to be brought to police departments to start collecting more in-depth data digitally to enhance better predictions. For example, variables such as on-duty officers pairings, officer performance in the academy, number of internal complaints, amount of inter-police department cases, time check-out sheets, etc., would greatly benefit researchers in tracking the contagious effect between officers regarding misconduct and violence.

Second, more research must be done on different police departments, not just only the Invisible Institute. Other papers mentioned in Chapter 2: Literature Review, used different datasets. More work could be done to see whether the proposed algorithms are transferable.

Last, more research should be done using different methods to predict excessive force. Many point-processes models, such as the Hawkes Process model based on the Poisson process, could be used in place of our algorithms to determine temporal events. We could use these models to estimate parameters such as reproduction numbers (i.e., the average number of excessive force cases arising from one police officer). If more variables are available, some graph-based techniques, such as graph neural networks, could be used to estimate the transmission of violence between police officers, potentially giving us more results.

To conclude, we believe that better data collection methods and more attention to the topic would be beneficial in preventing excessive use of force by the police.

Bibliography

[1] E. Hill, A. Tiefenthäler, C. Triebert, D. Jordan, H. Willis, and R. Stein, "How george floyd was killed in police custody," *The New York Times*, 2020.

[2] T. N. Y. Times, "Police brutality, misconduct and shootings," *The New York Times*, 2023.

[3] J. S. Motley RO Jr, "Police use of force by ethnicity, sex, and socioeconomic class.," *Journal of the Society for Social Work and Research*, vol. 81, no. 5, pp. 1222–1228, 2018.

[4] G. Slutkin, C. Ransford, and R. B. Decker, *Cure Violence: Treating Violence As a Contagious Disease*, pp. 43–56. Cham: Springer International Publishing, 2015.

[5] A. V. Papachristos, D. M. Hureau, and A. A. Braga, "The corner and the crew: The influence of geography and social networks on gang violence," *American Sociological Review*, vol. 78, no. 3, pp. 417–447, 2013.

[6] A. Papachristos, A. Braga, and D. Hureau, "Social networks and the risk of gunshot injury," *SSRN Electronic Journal*, 01 2011.

[7] R. Bond and B. Bushman, "The contagious spread of violence among us adolescents through social networks," *American Journal of Public Health*, vol. 107, pp. 288–294, 02 2017.

[8] M. Ouellet, S. Hashimi, J. Gravel, and A. V. Papachristos, "Network exposure and excessive use of force," *Criminology & Public Policy*, vol. 18, no. 3, pp. 675–704, 2019.

[9] G. Wood, D. Roithmayr, and A. V. Papachristos, "The network structure of police misconduct," *Socius*, vol. 5, p. 2378023119879798, 2019.

[10] A. Chalfin and J. Kaplan, "How many complaints against police officers can be abated by incapacitating a few "bad apples?"," *Criminology & Public Policy*, vol. 20, no. 2, pp. 351–370, 2021.

[11] M. Boss and J. Kleinert, "Motivational contagion during exercise and the role of interpersonal relationships: An experimental study," *PsyCh Journal*, vol. 10, 09 2020.

[12] D. Weisburd, R. Greenspan, E. Hamilton, H. Williams, and K. Bryant, "Police attitudes toward abuse of authority: Findings from a national study," *National Institute of Justice*, 5 2000.

[13] C. W. Telep, "The impact of higher education on police officer attitudes toward abuse of authority," *Journal of Criminal Justice Education*, vol. 22, no. 3, pp. 392–419, 2011.

[14] F. Sharara, E. Wool, G. Bertolacci, N. Weaver, S. Balassyano, A. Watson, I. Dippenaar, M. Cunningham, J. Fuller, L. Marczak, C. Johnson, G. Roth, A. Abedi, J. Acuna, S. Advani, P. Agasthi, F. Alahdab, Z. Al-Aly, C. Antony, and M. Naghavi, "Fatal police violence by race and state in the usa, 1980-2019: a network meta-regression," *The Lancet*, vol. 398, pp. 1239–1255, 10 2021.

[15] S. Alang, D. McAlpine, E. McCreedy, and R. Hardeman, "Police brutality and black health: Setting the agenda for public health scholars," *American Journal of Public Health*, vol. 107, no. 5, pp. 662–665, 2017. PMID: 28323470.

[16] H. L. Cooper, "War on drugs policing and police brutality," *Substance Use & Misuse*, vol. 50, no. 8-9, pp. 1188–1194, 2015. PMID: 25775311.

[17] P. Lyle and A. M. Esmail, "Sworn to protect: Police brutality – a dilemma for america's police," *Race, Gender & Class*, vol. 23, no. 3-4, pp. 155–185, 2016.

[18] K. M. Lersch and T. Mieczkowski, "Violent police behavior: Past, present, and future research directions," *Aggression and Violent Behavior*, vol. 10, no. 5, pp. 552–568, 2005.

[19] D. Eitle, S. J. D'Alessio, and L. Stolzenberg, "The effect of organizational and environmental factors on police misconduct," *Police Quarterly*, vol. 17, no. 2, pp. 103–126, 2014.

[20] J. Greene, A. Piquero, M. Hickman, and L. Brian, "Police integrity and accountability in philadelphia: Predicting and assessing police misconduct," *U.S. Department of Justice Office of Justice Programs*, 12 2004.

[21] D. Griswold, "Complaints against the police: Predicting dispositions," *Journal of Criminal Justice*, pp. 215–221, 1994.

[22] L. Zhao and A. V. Papachristos, "Network position and police who shoot," *The ANNALS of the American Academy of Political and Social Science*, vol. 687, no. 1, pp. 89–112, 2020.

[23] M. Sierra-Arévalo and A. Papachristos, "Bad apples and incredible certitude," *Criminology & Public Policy*, vol. 20, no. 2, pp. 371–381, 2021.

[24] C. Simpson and D. Kirk, "Is police misconduct contagious? non-trivial null findings from dallas, texas," *Journal of Quantitative Criminology*, pp. 1–39, 01 2022.

[25] "Invisible institue, howpublished = https://invisible.institute/, note = Accessed: 2023-04-17."

[26] G. Wood, T. R. Tyler, and A. V. Papachristos, "Procedural justice training reduces police use of force and complaints against officers," *Proceedings of the National Academy of Sciences*, vol. 117, pp. 9815–9821, May 2020.

Appendix A

Code and Figures

A.1 Code

A.1.1 Alpha Optimization Code

https://colab.research.google.com/drive/18tC2oznPrCKSD1PXfUzUQ1jzYO6JFKQk?usp=sharing

A.1.2 Beta Optimization Code

https://colab.research.google.com/drive/1WF6cH31hYpu_D7VXVTFy8jY2tJjr4cI8?usp=sharing

A.1.3 Visualization Code

https://colab.research.google.com/drive/1f7q5DOEqUuA4MH_m4lfTvWPIu3xARQht#scrollTo=Z-fdwnfLyKAS

A.1.4 GitHub Repository

For preprocessing data code, please visit:

https://github.com/Jeremyhudsonchan/Undergraduate-Thesis

A.1.5 Combined Results

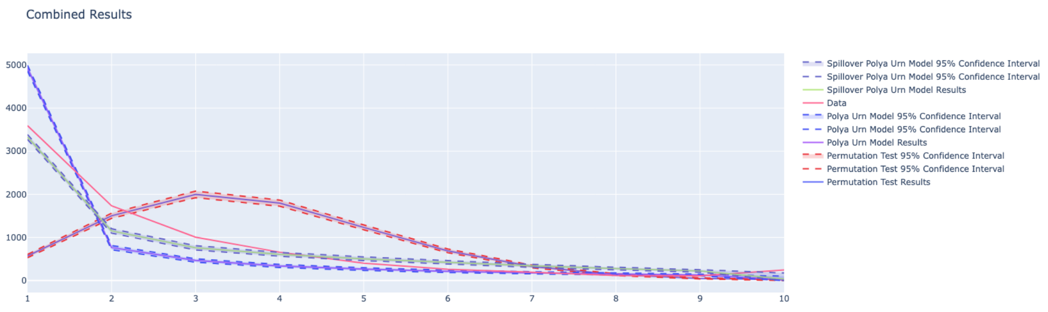

Figure A.1: Spillover Pólya Urn Model Results

Algorithm	Mean Squared Error
Resampling	1241392.95
Pólya Urn	290696.78
Spillover Pólya Urn	60710.71

Table A.1: Mean Squared Error of Each Algorithm Compare to Data