2026 American Control Conference (ACC)
May 27-29, 2026. New Orleans, Louisiana, USA

Mixed Integer Optimization of Hawkes Process Networks with Applications to Reducing Police Use-Of-Force*

Youness Diouane[1] and George Mohler[2]

***Abstract*— We propose a mixed-integer linear programming (MILP) framework to reduce police use-of-force incidents, designed to complement existing prevention tools such as early intervention systems (EIS). We model use-of-force incidents using a multivariate Hawkes process and optimize officer pairing networks through MILP to minimize the expected number of use-of-force events. In addition, we introduce a new risk metric, the reproduction number θ, which quantifies how likely an officer's use-of-force incident is to trigger subsequent incidents, either by the same officer or by colleagues nearby in the network. Our results indicate that reducing use-of-force requires pairing high-θ officers with low-θ officers. By preventing clusters of high-risk officers, the optimized network suppresses social contagion of risky behaviors, disrupts escalation chains, and reduces the forecasted frequency of use-of-force incidents.**

I. Introduction

Police use-of-force refers to the authority granted by law enforcement officers to employ physical coercion to ensure public safety, enforce laws, and protect themselves from harm. This authority is a critical component of policing, but it intersects with fundamental issues regarding civil rights, public trust, and community relations [1]. Police officers are tasked with making a variety of decisions, and their ability to use force is one of the defining characteristics of police work [2].

Past research on use-of-force risk factors has focused on race [3]–[5], involvement in violent crimes [6] and situational factors that can lead to the use of lethal force [7]. Other research has focused on officer characteristics; for example, officers with negative marks on their record are three times more likely to use lethal force in the course of their duty [8].

In practice, most police departments rely on early intervention systems (EIS) that flag officers when they cross simple thresholds, e.g., a set of number of complaints, use-of-force incidents, or disciplinary records [9]–[11]. Thresholds-based systems face two main limitations. First, they rely on raw observed counts of events, rather than estimated risk from a statistical model. Second, they fail to capture the ways in which one officer's actions may influence the behavior of other officers in their social network. This second issue is important, as recent research has shown that officer shootings, misconduct and excessive use-of-force exhibit network effects, where officers are at greater risk of being involved in these incidents when they socialize with officers who have a history of misconduct and complaints [12]–[15]. Patterns of excessive use-of-force can dynamically emerge from local interactions between individuals that aggregate to form global patterns of escalation and contagion [16].

In this paper, we introduce a method for optimizing police officer networks to reduce social contagion of use-of-force. In particular, we combine multivariate Hawkes processes (MHP's) with mixed integer linear programming (MILP) to design officer-assignment networks that minimize the expected number of use-of-force events. We leverage Hawkes processes to model how force incidents propagate through officer networks, and formulate officer assignment as an MILP that directly optimizes the network structure to minimize the expected number of incidents. We validate our approach on five U.S police department datasets, and demonstrate that it consistently reduces the expected number of incidents, outperforming standard baselines such as random permutation and rotation policies. The study also shows that the optimized networks systematically pair high-risk officers (officers that are more likely to use force) with low-risk officers, which suggests that structural interventions at the assignment level can contain escalation and prevent clustering of risky behaviors.

The remainder of the paper is organized as follows: Section II details the Hawkes process model and the MILP formulation. Section III introduces the datasets used for validation, and Section IV presents the results. Section V provides a discussion of the findings and their implications for policing and risk prevention, as well as the main limitations of the proposed framework.

II. Method

Hawkes processes [17] are point processes that are widely applied to a variety of domains, including seismology [18], epidemic diseases [19], neural activity [20], finance [21]–[23] and criminal justice [24], [25]. Hawkes processes model contagion effects between events, where one event in the history of the process can increase the likelihood of a subsequent event occurring in the near future. In the present application, we use Hawkes processes to model how a use-of-force incident involving one officer can increase the risk of subsequent events in the officers' social network.

Past research on control of Hawkes processes has included optimizing when to post for smart broadcasting [26], optimal

*This work was supported in part by SCC-2125319 and AFOSR MURI grant FA9550-22-1-0380

[1]Youness Diouane is with the Department of Computer Science, Boston College, 245 Beacon Street, Chestnut Hill, MA 02467, USA diouane@bc.edu

[2]George Mohler is with the Department of Computer Science, Boston College, 245 Beacon Street, Chestnut Hill, MA 02467, USA mohlerg@bc.edu

posting for fake news mitigation [27], and minimizing invasive species impacts through optimizing interventions [28]. However, these interventions focused on adding intervention events to the process, rather than rewiring the network to optimize the spread of events. To our knowledge the present paper is the first to consider optimizing the network on which the Hawkes process is defined through MILP.

A. Multivariate Hawkes Processes

Consider a network with $i = 1 \dots n$ nodes, where at each node the conditional intensity is given by,

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{n} \sum_{t_j^k < t} \alpha_{ij}\theta_j\omega_{ij}e^{-\omega_{ij}(t-t_j^k)}, \tag{1}$$

where μ_i determines the background intensity (assumed to be Poisson), and t_j^k is the timestamp of the k^{th} event at node j. The parameter α_{ij} is an indicator variable that takes the value 1 when i and j are connected in the network and 0 otherwise. In the present application, α determines the adjacency matrix of the police social network. For instance, when officer i and officer j are connected, $\alpha_{ij} = 1$. We also assume that each officer is connected to themselves, i.e., $\alpha_{ii} = 1$ (which models the fact that officers involved in use-of-force incidents are more likely to do so again in the future). The parameters θ_j represents the reproduction parameter, i.e., the expected number of events at node i directly triggered by a single event at node j when $\alpha_{ij} = 1$. The parameter ω_{ij} determines the timescale of secondary events at node i that are directly triggered by a primary event at node j.

B. Mixed integer optimization

Our approach to optimizing the Hawkes process network is to first derive its stationary distribution, and then to minimize the stationary distribution using mixed integer linear programming.

The stationary distribution of the Hawkes process in (1) is given by [29],

$$\begin{aligned}\bar{\lambda} &= \mathbb{E}[\lambda_i(t)] \\ &= \mu_i + \sum_{j=1}^{n} \mathbb{E}\left(\sum_{t_j^k<t} \alpha_{ij}\theta_j\omega_{ij}e^{-\omega_{ij}(t-t_j^k)}\right) \\ &= \mu_i + \sum_{j=1}^{n} \int_{-\infty}^{t} \alpha_{ij}\theta_j\omega_{ij}\bar{\lambda}_j e^{-\omega_{ij}(t-s)} ds \\ &= \mu_i + \sum_{j=1}^{n} \alpha_{ij}\theta_j\bar{\lambda}_j \left[e^{-\omega_{ij}(t-s)}\right]_{-\infty}^{t} \\ &= \mu_i + \sum_{j=1}^{n} \alpha_{ij}\theta_j\bar{\lambda}_j.\end{aligned} \tag{2}$$

Letting $g_{ij} = \alpha_{ij}\theta_j\bar{\lambda}_j$, we get

$$\begin{aligned} &\bar{\lambda}_i = \mu_i + \sum_{j=1}^{n} \alpha_{ij}\theta_j\bar{\lambda}_j \\ \Leftrightarrow\quad & \mu_i + \sum_{j=1}^{n} g_{ij} = \mu_i + \sum_{j=1}^{n} \alpha_{ij}\theta_j\left(\mu_j + \sum_{k=1}^{n} g_{jk}\right) \\ \Leftrightarrow\quad & \sum_{j=1}^{n} g_{ij} = \sum_{j=1}^{n} \alpha_{ij}\theta_j\left(\mu_j + \sum_{k=1}^{n} g_{jk}\right) \\ \Leftrightarrow\quad & g_{ij} = \alpha_{ij}\theta_j\left(\mu_j + \sum_{k=1}^{n} g_{jk}\right). \end{aligned} \tag{3}$$

Thus the stationary distribution is determined by the linear system (4):

$$-g_{ij} + \alpha_{ij}\theta_j\left(\mu_j + \sum_{k=1}^{n} g_{jk}\right) = 0. \tag{4}$$

Our goal is to minimize the total use of force in the system, $\sum_{i=1}^{n} \bar{\lambda}_i$. Since the background intensity μ_i is fixed, the minimization is equivalent to minimizing the term $\sum_{i=1}^{n} g_{ij}$. This leads to the following MILP problem:

$$\min_{\vec{\alpha}} 1^\top \vec{g}, \tag{5}$$

subject to $\alpha_{ij} = \{0,1\}$, $\alpha_{ii} = 1$, $\alpha = \alpha^\top$ ($\top$ denoting the transpose) and the linear constraints

$$\sum_{i=1}^{n} \alpha_{ij} = M+1 \tag{6}$$

$$\sum_{j=1}^{n} \alpha_{ij} = M+1 \tag{7}$$

$$-g_{ij} + \alpha_{ij}\theta_j\left(\mu_j + \sum_{k=1}^{n} g_{jk}\right) = 0 \tag{8}$$

Here each officer is connected to M other officers in the network. The $M+1$ means that in addition to M officers, each officer is considered connected to themselves. Note that M is a parameter that we fix and is not estimated from the data (see Algorithm 1). The optimizer we use to solve the above MILP is the IBM optimization software CPLEX 22.1.1.0 on an Intel(R) Core(TM) i9-7920X CPU @ 2.90GHz, 128GiB System memory machine.

Note that in (5), $1^\top$ denotes the row vector of ones and $\vec{g}$ is the vector obtained by stacking the elements g_{ij}. Therefore, the product $1^\top\vec{g}$ represents the sum of all elements of $\vec{g}$.

C. Maximum likelihood

The above optimization assumes that, other than $\vec{\alpha}$, the Hawkes process parameters are known. Our approach is to first use empirical data from several U.S. police departments to determine a historical value for $\vec{\alpha}$, when multiple officers are involved in the same incident. We then determine the parameters of the Hawkes process through maximum likelihood estimation. Finally, using those fixed parameters (but allowing $\vec{\alpha}$ to vary), we determine the amount by which forecasted use-of-force can be reduced.

Given a point process conditional intensity $\lambda(t)$, the log-likelihood function L_t with respect to the event times $t_1, t_2, \ldots, t_n$, is given by

$$L = \sum_{k=1}^{n} \log(\lambda(t_k)) - \int_0^T \lambda(s)ds, \tag{9}$$

where T is the observation time window of the data. To evaluate (9), the integral $\int_0^t \lambda(s)ds$ must be analytically determined or approximated. Let N_i be the number of events at node i. Therefore, the log-likelihood is given by:

$$\begin{aligned} L &= \sum_{i=1}^{n}\left(\sum_{k=1}^{N^i} \log \lambda_i(t_i^k) - \int_0^T \lambda_i(s)ds\right) \\ &= \sum_{i=1}^{n}\sum_{k=1}^{N^i} \log \lambda_i(t_i^k) - \sum_{i=1}^{n} \mu_i T \qquad (10) \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t_j^k<t}^{N^j} \alpha_{ij}\theta_j \left[e^{-\omega_{ij}(s-t_j^k)}\right]_{t_j^k}^{T} \end{aligned}$$

We use PyTorch to optimize (10). After fitting the model to historical use-of-force data, we use those parameters to solve the MILP. The resulting optimal adjacency matrix (rewired network) α is then used to simulate a forecasted Hawkes process. The resulting volume of use-of-force incidents is then compared to simulated incidents using alternative specifications for $\vec{\alpha}$. We depict this procedure in Algorithm 1.

Algorithm 1 Police Network Optimization

1: Define network parameters μ, ω, θ, and the observation time window T.
2: Construct the officers pairing α^0 from historical police data.
3: Fit μ, ω, θ, T from police data using MLE:

$$\begin{aligned} \underset{\hat{\mu},\hat{\theta},\hat{\omega}}{argmax} L = \sum_{i=1}^{n}\sum_{k=1}^{N^i} \log \lambda_i(t_i^k) - \sum_{i=1}^{n} \mu_i T \\ - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{N^i} \alpha_{ij}^0 \theta_j \left(1 - e^{-\omega_{ij}\left(T - t_j^k\right)}\right). \end{aligned}$$

4: Using the estimated parameters $\hat{\mu}$, $\hat{\theta}$ and $\hat{\omega}$ Solve the MILP $\underset{\vec{\alpha}}{min} 1^\top \vec{g}$, subject to

$$\begin{aligned} &\sum_{i=1}^{n} \alpha_{ij} = M+1, \\ &\sum_{j=1}^{n} \alpha_{ij} = M+1, \\ &-g_{ij} + \alpha_{ij}\hat{\theta}_j\left(\hat{\mu}_j + \sum_{k=1}^{n} g_{jk}\right) = 0 \end{aligned}$$

Note that the parameter M is not estimated from the data here or in preceding discussion.
5: Simulate an optimal HP using the solution of the MILP α
6: Use Hawkes process simulation to evaluate the reduction in forecasted use-of-force.

III. Experiments

Before introducing our experiments, we provide definitions of several policing terms that will be used throughout this section:

1) *Beat*: A geographic area assigned to a police officer or patrol unit as part of their routine patrol responsibilities.
2) *Incident ID*: A unique identifier assigned to a specific police incident or report in the department's database.
3) *Badge number*: A unique identification number assigned to each police officer, used to distinguish officers in official records.
4) *Internal Affairs case number*: A unique identifier assigned to an investigation conducted by the Internal Affairs division regarding complaints, misconduct, or disciplinary matters involving police officers.

We apply our police network optimization algorithm on five different datasets:

- Chicago [1]: The dataset comprises complaints filed against Chicago Police Department officers, including those lodged by civilians, which make up the majority, as well as those filed by fellow officers. We limited our analysis to the period from January 1, 2022 to December 31, 2022, resulting in a network of 187 unique officers and a total of 1371 incidents. Two officers are considered connected if they share the same incident ID, beat, and incident time. Officers are distinguished by their full names.
- Seattle [2]: The data is collected from public records documenting use-of-force incidents involving sworn law enforcement officers of the Seattle Police Department. The period covered is from January 1 to October 31, 2023. As a result, the police network for this timeframe consists of 372 unique officers, and 1025 incidents overall. Two officers are considered connected if they share the same incident ID, beat, and time. Officers are distinguished by their badge number.
- Massachusetts[3]: The data is obtained from the Massachusetts POST Commission Law Enforcement Agency which publishes statewide records of complaints and disciplinary actions. The period examined is from January 1, 2024 to November 13, 2024. During this period, the police network includes 144 unique officers and 332 incidents. Two officers are considered connected if they are involved in incidents occurring at the same time and share similar Internal Affairs case

[1] https://github.com/invinst/chicago-police-data/tree/master/data/unified_data/complaints

[2] https://data.seattle.gov/Public-Safety/Use-Of-Force/ppi5-g2bj/data_preview

[3] https://mapostcommission.gov/discipline-status-records/disciplinary-records/

numbers. Officers are distinguished by their full names.

- Phoenix [4]: This dataset contains use-of-force incidents from January 2018 onward, including officer demographics, badge numbers, incident times, and counties. For our study we restrict the analysis to the period from January 1, 2024 to December 31, 2024. During this window we focus on 10 beats: 912, 825, 423, 422, 831, 425, 711, 932, 815, and 832. The data covers 224 officers and 331 incidents. Two officers are considered connected if they share the same incident time, beat, and incident number. Officers are distinguished by their badge numbers.
- Indianapolis [5]: The data comes from public records of use-of-force incidents provided by the Indianapolis Metropolitan Police Department (IMPD) use-of-force database. The period examined is from January 1, 2024 to December 31, 2024, and covers four patrol districts: East, Southwest, Northwest and Downtown. The resulting network includes 479 police officers and a total of 9875 incidents. Two officers are considered connected if they share the same incident time and incident ID. Officers are distinguished by their badge numbers.

To verify that our algorithm can determine the optimal network structure, we first simulate a small synthetic network that consists of six nodes ($n = 6$), where each officer is connected to two officers ($M = 2$). The network parameters are selected as:

$$
\begin{aligned}
\theta &= (0.1, 0.15, 0.15, 0.85, 0.8, 0.9) \\
\mu &= (0.1, 0.2, 0.1, 0.15, 0.2, 0.18) \\
\alpha^0 &= \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \\
\omega &= \begin{pmatrix} 0.8 & 0.2 & 0.3 & 0.5 & 0.2 & 0.3 \\ 0.3 & 0.2 & 0.5 & 0.5 & 0.6 & 0.2 \\ 0.9 & 0.2 & 0.3 & 0.8 & 0.6 & 0.7 \\ 0.9 & 0.2 & 0.4 & 0.9 & 0.6 & 0.5 \\ 0.8 & 0.2 & 0.4 & 0.9 & 0.5 & 0.4 \\ 0.8 & 0.2 & 0.4 & 0.3 & 0.2 & 0.1 \end{pmatrix}
\end{aligned}
\tag{11}
$$

We then exhaustively enumerate all possible 6×6 adjacency matrices, where $\sum_i \alpha_{ij} = 3$ and $\sum_j \alpha_{ij} = 3$, and verify that the adjacency matrix that minimizes the objective function $\sum_i g_{ij}$ matches the one determined by CPLEX and the formulation in Equations 5-8.

In this case there are exactly 70 possible adjacency matrices. This follows from the fact that the off-diagonal structure must form a 2-regular graph on six vertices, which can only be either a single 6-cycle or two disjoint 3-cycles. Fig. 1 illustrates the two topologies.

[4]https://www.phoenixopendata.com/dataset/ouof/resource/c79b2135-e936-439e-a8a3-79e61d4518d2

[5]https://data.indy.gov/datasets/IndyGIS::impd-use-of-force/about

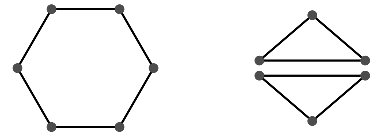

Fig. 1. The two possible topologies on six nodes: (left) a single 6-cycle and (right) two disjoint 3-cycles. These are the only unlabeled off-diagonal patterns; permuting labels yields the 70 feasible adjacency matrices (with $M = 2$ and $\alpha_{ii} = 1$)

Let m_n be the n-vertices cycle multiplicity, r the number of cycles, and s_i the number of vertices in cycle i. The number of permutations we can make is

$$\frac{n!}{2^r \prod_i m_i! \prod_j s_j} \tag{12}$$

For the two possible topologies in Fig. 1, this yields

$$\frac{6!}{2^1 \cdot 1! \cdot 6} + \frac{6!}{2^2 \cdot 2! \cdot 3 \cdot 3} = 70 \tag{13}$$

After enumerating all 70 possibilities, we confirmed that the adjacency matrix minimizing $\sum_i g_{ij}$ is identical to the one obtained by the CPLEX optimizer [6]. The optimal adjacency matrix is:

$$\alpha = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{14}$$

From (14), we observe that the optimizer tends to pair officers with high values of θ to officers with low value of θ. A more detailed analysis of assortativity patterns is provided in Subsection IV-A.

Finally, we validate our model parameter estimation by plotting the sampling distributions of the maximum likelihood estimates for the model parameters θ, μ, and ω for a small synthetic network of four officers across 1000 simulations, along with the true parameter values shown as dashed blue lines in Fig. 2. Notably, for all parameters and across all nodes, the true values are contained within the corresponding sampling distribution.

IV. Results

We next apply the network optimization algorithm to the 5 police department datasets. We compare our proposed MILP-based optimization algorithm to three baseline approaches: ”greedy influence heuristic”, ”random officer reassignment” and ”rotation based policies”. The greedy influence heuristic algorithm repeatedly removes the officer-to-officer edge that is driving the largest amount of self-excitation in the fitted Hawkes model. On the other hand, random officer reassignment algorithm produces a shuffled version of the

[6]The code for this verification as well as the results for all the experiments is available at https://github.com/younesszs/Police-use-of-force.git

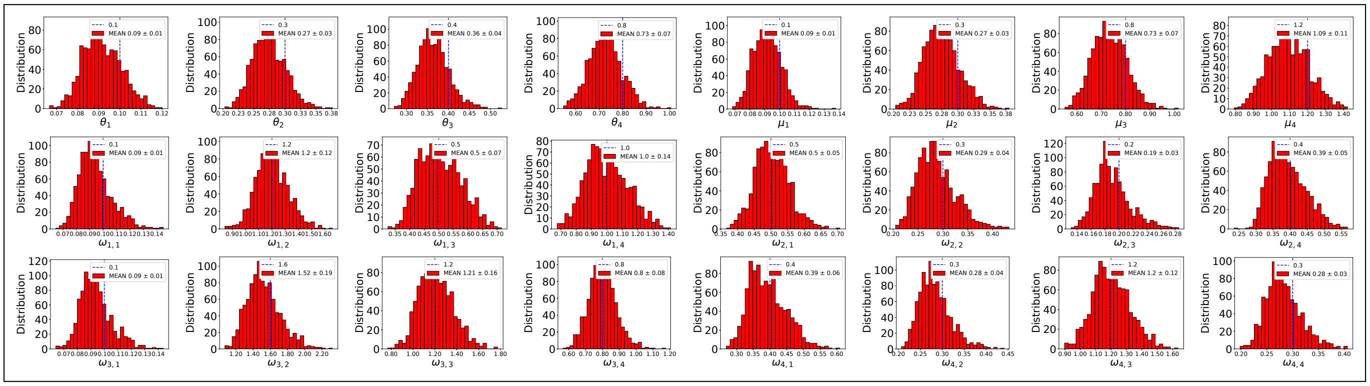

Fig. 2. Sampling distributions of the maximum likelihood estimates of the Hawkes process parameters θ, μ and ω across 1000 simulations of a synthetic network. Each histogram shows the empirical distribution of the estimated parameter values; the dashed blue vertical line marks the corresponding ground-truth parameter used in data generation. The legend reports the true value alongside the empirical mean and standard deviation of the estimates. The synthetic network is described by $n=4$, $T=100$, $\theta=(0.1,0.3,0.4,0.8)$, $\mu=(0.1,0.3,0.8,1.2)$, $\omega=\begin{pmatrix}0.1 & 1.2 & 0.5 & 1\\ 0.5 & 0.3 & 0.2 & 0.4\\ 0.1 & 1.6 & 1.2 & 0.8\\ 0.4 & 0.3 & 1.2 & 0.3\end{pmatrix}$, and $\alpha^0=\begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 1 & 1\\ 1 & 1 & 1 & 0\\ 1 & 1 & 0 & 1\end{pmatrix}$.

MHP where officers are randomly paired to M other officers. Finally, the rotation-based policy models the effect of a rotation schedule on the MHP. The rotation schedule randomly reassigns each officer's connections every Δt time units.

We present the results in Table I, which reports the mean number of forecasted use-of-force incidents across the different police networks (Chicago, Seattle, Massachusetts, Phoenix, and Indianapolis). In all experiments, we set $M=8$, applied the rotation policy with a 90-day period, and used a CPLEX time limit of 8000 seconds. The table also includes a baseline comparison between our proposed optimization framework and the alternative approaches described previously. Across all experiments, we observe that the proposed method consistently outperforms the baselines.

A. *Assortativity analysis*

Assortativity is a measure used to determine whether nodes in a network are preferentially connected to other nodes that are similar (or dissimilar) to themselves with respect to some attribute, e.g., node degree. Assortativity essentially accounts for the correlation between nodal properties of connected pairs in the network. If nodes with similar properties are more likely to be connected, the network is said to be assortative; conversely, if nodes with dissimilar properties are more likely to be connected, the network is said to be disassortative. In our case, the node attribute we focus on is the reproduction number θ_j, since it is a measure of how an officer's use-of-force incident increases the risk of future incidents, either by themselves or by others nearby them in the network. We examine the assortativity of θ_j by calculating the assortativity coefficient r for each police network before and after optimization. The coefficient r represent the Pearson correlation between each (θ_i,θ_j) pair. It ranges between -1 and 1. Values where $r>0$ indicate assortative mixing, $r<0$ signals disassortative mixing, and $r=0$ (or close to zero) signals neutral mixing.

In Table II we display r for each network after optimization, and in Fig. 3 we plot five graphs corresponding to each optimized police network, highlighting a randomly selected node and its connections. Note that in our setting, high-θ officers correspond to officers with $\theta \geq 0.5$ whereas low-θ officers are the ones with $\theta < 0.5$.

From Table II and Fig. 3, we observe that after optimization and in all experiments, the coefficient r is strictly negative, indicating disassortative mixing, which means that nodes with low values θ tend to connect with nodes that have high values θ and vice versa. This result indicates that, in order to decrease use-of-force, police pairing should aim to connect officers with a high reproduction number to officers with a low reproduction number. This result matches past research on network percolation showing that assortative networks facilitate network spreading more than their dissassortative counterparts [30]. Furthermore, assortative networks are more susceptible to epidemic outbreaks, whereas disassortative networks display a higher epidemiological threshold and are therefore easier to immunize [31].

V. Conclusion

In this work, we introduced an MILP-based optimization framework for reducing the expected number of police use-of-force incidents, modeled through a multivariate Hawkes process. Using five different datasets across the U.S, i.e., Chicago, Seattle, Massachusetts, Indianapolis, and Phoenix, we demonstrated that our optimization consistently outperforms baseline methods such as random permutation and rotation policies.

A key finding of our approach is that the optimized networks exhibit dissassortative mixing, where officers with high reproduction number θ, are paired with officers who have low reproduction number. The parameter θ_j quantifies the propagation potential of officer j's incidents, i.e., how likely a use-of-force event involving j increases the likelihood of subsequent events by j or their close colleagues in the near future. Conceptually θ is analogous to the reproduction number R_0 in epidemiology, which ultimately

TABLE I

AVERAGE NUMBER OF FORECASTED POLICE USE-OF-FORCE INCIDENTS USING DIFFERENT APPROACHES FOR THE DIFFERENT POLICE NETWORKS I.E., CHICAGO, SEATTLE, MASSACHUSETTS, PHOENIX AND INDIANAPOLIS OVER 1000 SIMULATIONS. IN ALL EXPERIMENTS WE SET $M = 8$. THE ROTATION POLICY WAS APPLIED WITH A ROTATION PERIOD OF 90 DAYS, AND THE OPTIMIZER WAS RUN WITH A TIME LIMIT OF 8000 SECONDS.

Approach	Chicago	Seattle	Massachusetts	Phoenix	Indianapolis
Hawkes MILP	**590.405**	**474.193**	**508.467**	**320.333**	**704.552**
Greedy heuristic	823.488	763.452	747.689	840.354	898.228
Random reassignment	830.892	767.516	754.573	847.128	902.250
Rotation policy	802.172	738.769	727.809	816.440	870.371

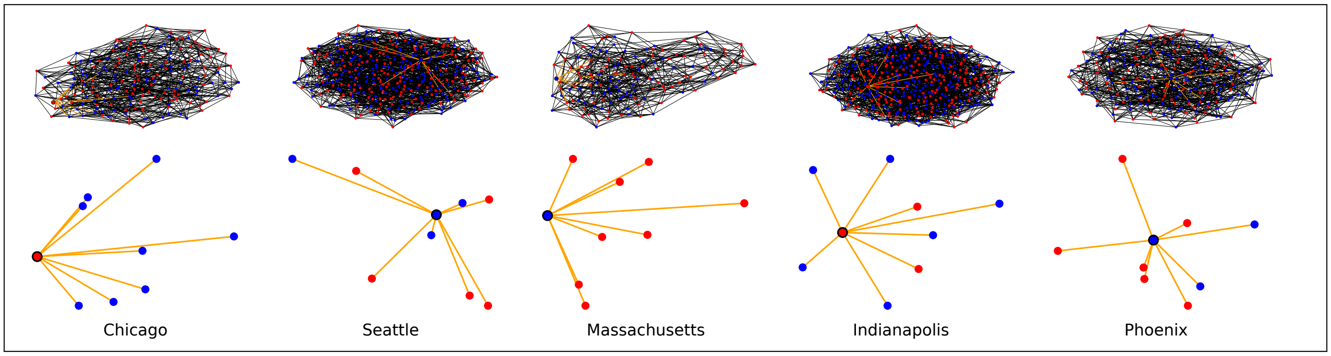

Fig. 3. Optimized network ($M = 8$) configurations for the five police departments (left to right: Chicago, Seattle, Massachusetts, Indianapolis, Phoenix). The top row shows the full optimized networks, where a randomly selected officer is highlighted along with their connections (edges in orange). The bottom row provides a zoomed-in view of the same officer and their immediate neighborhood. Nodes are colored according to their reproduction number θ i,e., blue for $\theta < 0.5$ (low-θ officers) and red for $\theta \geq 0.5$ (high-θ officers). Across departments, these neighborhoods indicate a disassortative pattern in which high-θ officers tend to be linked to low-θ officers and vice versa.

TABLE II

ASSORTATIVITY COEFFICIENT r FOR THE FIVE POLICE NETWORKS AFTER IMPLEMENTING THE MILP OPTIMIZATION. OUR OPTIMIZER WAS SET TO A TIME LIMIT OF 8000 SECONDS.

Network	Assortativity coefficient
Chicago	-0.72
Seattle	-0.29
Massachusetts	-0.72
Indianapolis	-0.16
Phoenix	-0.56

determines how fast and to what extent infectious diseases spread through a population [32].

The fact that our optimized networks consistently link high-θ officers with low-θ officers highlights the potential for θ to be used as a risk metric for guiding officer pairing and assignment. By contrast, in most major American cities, police officers are paired and assigned to neighborhoods based on seniority. Such an experience-based approach may lead to misalignment between officers and the neighborhoods where they are assigned [33]; other factors such as race and age also influence pairing and assignment decisions [12], but to our knowledge use-of-force risk is not currently a factor used to make pairing decisions.

We want to stress that our framework is designed to be integrated into existing early-intervention systems, rather than replacing them. By incorporating θ as an additional indicator, departments could prevent high risk officers from being paired together. In practice, θ can be estimated though a multivariate Hawkes process, as we do in this study. However, for departments without in-house modeling capacity, other risk scores such as those from early-intervention systems, or count-based metrics, could be used instead.

Our study suggests two practical pathways for implementation. The first is to use the full MILP framework, which directly generates the optimized network structure and prescribes the assignment configuration that minimizes the expected number of use-of-force incidents. While effective, this approach can be computationally expensive (see limitations below) and may be logistically challenging for some departments. As a simpler alternative, agencies could adopt a heuristic strategy. In this approach, departments would estimate θ for each officer using a multivariate Hawkes, and then enforce disassortative pairing by systematically matching high-θ officers with their low-θ counterparts.

One limitation of our approach is its scalability with the number of nodes n. The complexity of the problem in Equations 5-8 is $O(n^3)$. For solvers such as CPLEX that rely on branch and bound [34], complexity grows exponentially with respect to the decision variable dimensions [35]. Given the dense constraint matrix of our problem, and the exponential runtime of CPLEX [36], we imposed CPLEX limit of 8000 seconds. Time limited runs are standard practice in mixed integer programming [37], [38], but we note that, for larger networks, runtime may be a hurdle to finding optimal police networks.

A second limitation of the present study is that we use *forecasted* use-of-force to assess the performance of our algorithm. While contagion effects in police networks have been observed in retrospective analyses, to our knowledge attempts have not yet been made to rewire real police networks in observational studies to determine if the risk of use-of-force and misconduct can be lowered. Furthermore, an officers individual risk, parameterized by μ and θ in the present study, may change over time as the officer gains experience. Future iterations of the present model might allow for time dynamic parameters, e.g. $\theta(t)$, and in that case dynamic optimization would be needed to control the system.

REFERENCES

[1] W. Terrill, "Deadly force: To shoot or not to shoot," *Criminology & Pub. Pol'y*, vol. 15, p. 491, 2016.

[2] E. Bittner, *The functions of the police in modern society: A review of background factors, current practices, and possible role models*. National Institute of Mental Health, Center for Studies of Crime and Delinquency, 1970.

[3] G. D. Robin, "Justifiable homicide by police officers," *The Journal of Criminal Law, Criminology, and Police Science*, vol. 54, no. 2, pp. 225–231, 1963.

[4] P. Takagi, "A garrison state in 'democratic'society," *Crime and Social Justice*, pp. 205–219, 1981.

[5] A. L. Kobler, "Figures (and perhaps some facts) on police killing of civilians in the united states, 1965–1969," *Journal of Social Issues*, vol. 31, no. 1, pp. 185–191, 1975.

[6] L. W. Sherman, "Execution without trial: Police homicide and the constitution," *Vand. L. Rev.*, vol. 33, p. 71, 1980.

[7] M. D. White, "Identifying situational predictors of police shootings using multivariate analysis," *Policing: an international journal of police strategies & management*, vol. 25, no. 4, pp. 726–751, 2002.

[8] G. Ridgeway, "Officer risk factors associated with police shootings: a matched case–control study," *Statistics and Public Policy*, vol. 3, no. 1, pp. 1–6, 2016.

[9] S. Walker, *Early intervention systems for law enforcement agencies: A planning and management guide*. US Department of Justice, Office of Community Oriented Policing Services, 2003.

[10] M. D. Fan, "Smarter early intervention systems for police in an era of pervasive recording," *U. Ill. L. Rev.*, p. 1705, 2018.

[11] C. M. Katz, H. Cheon, K. Freemon, and D. Wallace, "Evaluating the effectiveness of a police early intervention system: from the predictive validity of officer identification to the impact of intervention," *Police Quarterly*, p. 10986111251353487, 2024.

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

[13] M. Ouellet, S. Hashimi, J. Gravel, and A. V. Papachristos, "Network exposure and excessive use of force: Investigating the social transmission of police misconduct," *Criminology & Public Policy*, vol. 18, no. 3, pp. 675–704, 2019.

[14] 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.

[15] R. Pandey, J. Carter, J. Hill, and G. Mohler, "Rewiring police officer training networks to reduce forecasted use of force," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4696–4704, 2023.

[16] D. Roithmayr, "The dynamics of excessive force," *U. Chi. Legal F.*, p. 407, 2016.

[17] A. G. Hawkes, "Spectra of some self-exciting and mutually exciting point processes," *Biometrika*, vol. 58, no. 1, pp. 83–90, 1971.

[18] A. G. Hawkes, "Cluster models for earthquakes-regional comparisons," *Bull. Int. Stat. Inst.*, vol. 45, no. 3, pp. 454–461, 1973.

[19] Y. Diouane, F. Schoenberg, and G. Mohler, "Accurate estimation of cross-excitation in multivariate hawkes process models of infectious diseases," in *2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–8, IEEE, 2024.

[20] S. A. Pasha and V. Solo, "A nonlinear hawkes model using laguerre orthogonal polynomials," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, pp. 3477–3482, IEEE, 2024.

[21] Y. Aït-Sahalia, J. Cacho-Diaz, and R. J. Laeven, "Modeling financial contagion using mutually exciting jump processes," *Journal of Financial Economics*, vol. 117, no. 3, pp. 585–606, 2015.

[22] B. Gašperov and Z. Kostanjčar, "Deep reinforcement learning for market making under a hawkes process-based limit order book model," *IEEE control systems letters*, vol. 6, pp. 2485–2490, 2022.

[23] R. Luo and V. Krishnamurthy, "Detecting structural shifts in multivariate hawkes processes with fréchet statistic," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, pp. 2423–2428, IEEE, 2024.

[24] J. Leverso, Y. Diouane, and G. Mohler, "Measuring online–offline spillover of gang violence using bivariate hawkes processes," *Journal of Quantitative Criminology*, pp. 1–29, 2024.

[25] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-exciting point process modeling of crime," *Journal of the american statistical association*, vol. 106, no. 493, pp. 100–108, 2011.

[26] A. Zarezade, U. Upadhyay, H. R. Rabiee, and M. Gomez-Rodriguez, "Redqueen: An online algorithm for smart broadcasting in social networks," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 51–60, 2017.

[27] M. Farajtabar, J. Yang, X. Ye, H. Xu, R. Trivedi, E. Khalil, S. Li, L. Song, and H. Zha, "Fake news mitigation via point process based intervention," in *International conference on machine learning*, pp. 1097–1106, PMLR, 2017.

[28] A. Gupta, M. Farajtabar, B. Dilkina, and H. Zha, "Discrete interventions in hawkes processes with applications in invasive species management.," in *IJCAI*, pp. 3385–3392, 2018.

[29] X. Gao and L. Zhu, "Functional central limit theorems for stationary hawkes processes and application to infinite-server queues," *Queueing Systems*, vol. 90, pp. 161–206, 2018.

[30] M. E. Newman, "Mixing patterns in networks," *Physical review E*, vol. 67, no. 2, p. 026126, 2003.

[31] G. D'Agostino, A. Scala, V. Zlatić, and G. Caldarelli, "Robustness and assortativity for diffusion-like processes in scale-free networks," *Europhysics letters*, vol. 97, no. 6, p. 68006, 2012.

[32] P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen, "Complexity of the basic reproduction number (r0)," *Emerging infectious diseases*, vol. 25, no. 1, p. 1, 2019.

[33] B. Ba, P. Bayer, N. Rim, R. Rivera, and M. Sidibé, "Police officer assignment and neighborhood crime," tech. rep., National Bureau of Economic Research, 2021.

[34] A. H. Land and A. G. Doig, "An automatic method for solving discrete programming problems," in *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, pp. 105–132, Springer, 2009.

[35] L. Huang, X. Chen, W. Huo, J. Wang, F. Zhang, B. Bai, and L. Shi, "Branch and bound in mixed integer linear programming problems: A survey of techniques and trends," *arXiv preprint arXiv:2111.06257*, 2021.

[36] K. Ed and M. Newman Alexandra, "Practical guidelines for solving difficult linear programs," *Surveys in Operations Research and Management Science*, vol. 18, no. 1-2, pp. 1–17, 2013.

[37] P. Belotti, P. Bonami, M. Fischetti, A. Lodi, M. Monaci, A. Nogales-Gómez, and D. Salvagnin, "On handling indicator constraints in mixed integer programming," *Computational Optimization and Applications*, vol. 65, no. 3, pp. 545–566, 2016.

[38] J. Jablonskỳ *et al.*, "Benchmarks for current linear and mixed integer optimization solvers," *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 63, no. 6, pp. 1923–1928, 2015.