ELSEVIER

Towards social fairness in smart policing: Leveraging territorial, racial, and workload fairness in the police districting problem

Federico Liberatore [a,b,*], Miguel Camacho-Collados [c,d], Lara Quijano-Sánchez [e,b]

[a] School of Computer Science and Informatics, Cardiff University, Abacws Building, Cardiff, CF24 4AX, United Kingdom
[b] UC3M-Santander Big Data Institute, Charles III University of Madrid, Spain
[c] Department of Technological Innovation and Cybersecurity, Council of State, Madrid, Spain
[d] National Police, Ministry of the Interior, Spain
[e] School of Engineering, Autonomous University of Madrid, Madrid, Spain

ARTICLE INFO

ABSTRACT

Recent events (e.g., George Floyd protests) have shown the impact that inequality in policing can have on society. Thus, police operations should be planned and designed taking into account the interests of three main groups of directly affected stakeholders (i.e., general population, minorities, and police agents) to pursue fairness. Most models presented so far in the literature failed at this, optimizing cost efficiency or operational effectiveness instead while disregarding other social goals. In this paper, a Smart Policing model that produces operational patrolling districts and includes territorial, racial, and workload fairness criteria is proposed. The patrolling configurations are designed according to the territorial distribution of crime risk and population subgroups, while equalizing the total risk exposure across the districts, according to the preferences of a decision-maker. The model is formulated as a multi-objective mixed-integer program. Computational experiments on randomly generated data are used to empirically draw insights into the relationship between the fairness criteria considered. Finally, the model is tested and validated on a real-world dataset about the Central District of Madrid (Spain). Experiments show that the model identifies solutions that dominate the current patrolling configuration used.

1. Introduction

Predictive Policing is the use of police data to identify people, places and events with a high risk of crime. Crime predictions allow police efforts to be focused on the areas of greatest risk while dynamically adapting to the current criminal trends in a specific territory. Numerous pilot projects (Los Angeles [1], Chicago [2], and London [3], among others) have shown the capabilities of these methodologies and provided very good results. For example, PredPol, a computer tool used in Santa Cruz, California, uses historical data to point officers to 10–20 high-risk areas [4]. The result has been a reduction in the first year of 27% of burglaries and 11% of home burglaries. In a similar context, Baycik et al. [5] address the problem of allocating police resources to interdict criminal activities of different types; the model is defined within a Markov Decision Process framework.

However, crime prediction represents only the first step toward defining a police action protocol that fits the profile of each shift. *Smart Policing* [6] is the use of optimization techniques to determine the distribution of patrol agents in a district according to a number of objectives, thus, allowing for an improvement in the workload assignment between agents and an increase in the efficiency of actions [7–9]. The models are formulated as districting problems that have been widely and successfully applied in multiple contexts, especially in the public sector [10,11]. For an updated review of the most recent contributions in *Smart Policing* the reader is referred to Samanta et al. [12].

The patrol configurations obtained by models that only take into account the predicted distribution of crime tend to concentrate police resources in the areas with the highest crime rates [13]. However, high-crime locations are usually correlated with minority and disadvantaged populations [14]. Therefore, patrol operations designed from crime prediction data, despite being efficient and effective in controlling crime [15], could exacerbate the racial disparity in arrests and police actions [16,17] which, in turn, fuels the feeling of pressure in minority communities and, ultimately, can result in negative repercussions for both individuals and society as a whole (e.g., [18]). Nevertheless,

this issue cannot be solved by reducing police presence in high-crime areas, as it would deprive minority communities of much-needed police protection, resulting in discrimination and further victimization [19].

In recent years, researchers have addressed the issue of verifying the impact of considering racial equality criteria in Predictive Policing [20–22] and Smart Policing models [23,24], addressing a growing concern that is not restricted to the area of policing [25]. In particular, Liberatore et al. [24] segment the population into groups according to certain characteristics (e.g., origin, nationality) and consider the objective of equalizing police exposure among the different population groups. Their experimental results show that there is a trade-off between police efficiency and racial equality. However, a small reduction in police efficiency leads to a large increase in racial equality.

1.1. Summary and rationale

The rationale of this paper is to extend previous police districting models in the literature by balancing fairness criteria that represent the interests of three main groups of directly affected stakeholders: the general population, minorities, and police agents. In general, to counter crime, it is desirable that high-risk areas receive more patrolling time compared to low-risk areas. This is the reasoning behind the first criterion which, in the rest of the paper, is referred to as *territorial fairness*. However, as explained above, focusing the police efforts exclusively on high-risk areas could be counterproductive in the mid/long term. Second, the *racial fairness* criterion is introduced to ensure that the police exposition of the groups that comprise the population is proportioned.[1] However, as shown in the rest of this paper, designing patrolling districts while exclusively taking into account the above criteria results in configurations with high variability in terms of the exposure of the agents to the risk of crime which, in turn, produces an uneven distribution of the workload in terms of crime incidents that the agents need to prevent or respond. Additionally, greater equality in the workplace increases the satisfaction of the agents [26]. Therefore, the *workload fairness* criterion is considered to equalize the total exposition to the risk of crime across the districts.

Workload fairness has been previously considered in settings other than policing, such as in the Vehicle Routing Problem [27], assembly lines [28], and *seru* production systems [29,30]. In the context of police patrolling, multiple authors have provided different definitions of workload (please, refer to [12]). However, only Camacho-Collados and Liberatore [7],Liberatore and Camacho-Collados [9],Chen et al. [31] explicitly optimize for balance. All these authors define the workload as a linear combination of attributes, including total risk, total district area, and district diameter.

1.2. Contributions

Following from the above, this paper builds upon and expands on Liberatore et al. [24], forwarding and expanding the literature on Smart Policing and the Police Districting Problem in the following ways:

First, to the best of the authors' knowledge, this paper introduces the first police districting model that explicitly balances the interests of three groups of stakeholders (i.e., the general population, minorities, and police agents) by jointly considering territorial, racial, and workload fairness. Although the first two criteria have been previously introduced in [24] and workload has been formerly considered in police patrolling problems, this is the first time that the impact on the agents of a patrolling configuration is taken into account in terms of total exposure to the risk of crime. Thus, this is the first model that can be

used to identify solutions that are performant, socially equitable, and balanced in terms of workload.

Second, the model is used to perform a comprehensive empirical analysis that ascertains the impact of territorial, demographic, and crime characteristics on the trade-off between the fairness criteria. Thus, this research addresses an important gap in the literature.

Furthermore, two additional contributions lie in a novel dataset specifically generated to carry out the above analysis and a quantitative measure of trade-off strength, τ, that is completely agnostic to the present application context and can be applied to any bi-objective analysis.

Finally, the model is applied to a real-world case study which allows drawing insights into the usefulness and applicability of the model, as well as validating the outcomes of the empirical analysis. To this end, the solutions identified are compared to the districting configuration currently adopted by the Spanish National Police (SNP).

1.3. Outline

The rest of the paper is organized as follows. Section 2 presents in detail the problem and its formulation. The trade-off analysis is the topic of Section 3. The case study on the Central District of Madrid and the comparison with the solution currently in use by the SNP is illustrated in Section 4, while Section 5 concludes the paper with a summary of the main findings and guidelines for future research lines.

The paper includes two appendices. Appendix A describes the procedure used to generate the dataset while Appendix B presents the trade-off measure of strength, τ. Both are used in the trade-off analysis of Section 3.

2. Model

Let us consider a territory partitioned into areas (e.g., census or report districts). The territory is modeled as a graph, where the nodes represent the areas and two areas are connected by an edge if it is possible to move directly from one to the other. Also, the population in the territory is categorized into population groups, e.g., by race, ethnicity or origin.

The main decision in the problem involves assigning each area to a unique patrol district. The patrol districts must be geographically contiguous, their size is limited by a maximum diameter and they have a fixed time capacity (i.e., the length of the shift). The second main decision is the allocation of the time capacity available at each patrol districts to their assigned areas.

The crime risk and the population groups distributions on the territory are known. Therefore, each area is characterized by a crime level and by the size of each population group. These attributes are used to calculate (i) the patrol service goal for each area, (ii) the police contact goal for each population group, and (iii) the risk exposure goal for a district. Given a patrolling configuration (i.e., areas assignment to districts and time allocation to areas) it is possible to compute the following scores: (i) patrol service, i.e., the ratio of patrolling time received by an area to its goal; (ii) police contact, i.e., the ratio of the patrolling time received by a population group to its goal; (iii) risk exposure, i.e., the ratio of the total risk-weighted time spent by a district to its goal. Fairness is achieved by maximizing the weighted sum of the minimum of each score across the districts.

The rest of the section provides technical details of the methodology. A summary of the notation used is provided in Table 1. Readers only interested in the practical outcomes of this work may omit it.

---

[1] Please, note that although the criterion is called *racial fairness*, the model does not specify how the population groups are defined. For the sake of versatility, this choice is left to the decision-makers.

Table 1
Summary of notation: sets, indices, parameters, and variables.

Sets	
$G = (N, E)$	Graph.
N	Set of areas.
E	Set of edges connecting the areas.
K	Number of districts.
P	Set of population groups.
Indices	
$i, j \in N$	Areas.
$k = 1, \ldots, K$	District.
$p, q \in P$	Population group.
Parameters	
cap	Patrolling time capacity for each district.
$risk_i$	Crime risk associate to area i.
pop_{ip}	Size of population in area i belonging to group p.
$dist_{ij}$	Distance between areas i and j, expressed as the time necessary to go from i to j.
$diam$	Maximum district diameter.
$goal_i$	Patrol service goal for area i.
$goal_p$	Police contact goal for population group p.
$goal_k$	Risk exposure goal for district k.
w^T, w^R, w^W	Objective function weights for the territorial, racial, and workload fairness criteria.
Variables	
$assign_{ik}$	$\begin{cases} 1 & \text{if area } i \text{ is assigned to district } k, \\ 0 & \text{otherwise.} \end{cases}$
$time_{ik}$	Time allocated to area i in district k.
$source_{ik}$	$\begin{cases} 1 & \text{if area } i \text{ is the source of the flow of district } k \\ 0 & \text{otherwise} \end{cases}$
$flow_{ijk}$	Flow traversing edge (i, j) in district k.
$level_i$	Patrol service level of area i.
$level_p$	Police contact level of population group p.
$level_k$	Risk exposure level of district k.
W^T, W^R, W^W	Territorial, racial, and workload fairness criteria values.
W	Total fairness criterion value.

2.1. Model's parameters

Let $G = (N, E)$ be an undirected graph where N is the set of areas and E the set of edges connecting them. Also, K is the number of districts (indexed by k) and P is the set of population groups (indexed by p).

Districts have a constant time capacity, cap, while each area $i \in N$ is characterized by:

- a crime risk $risk_i$, representing the severity of the crimes that are expected to be committed in the area.
- a population distribution pop_{ip} representing the size of the population of area i that belongs to group p.

The edges have an associated distance, $dist_{ij}$. Moreover, the matrix $dist_{ij}^{SP}$ is the all-pair shortest-path distance matrix, computed from the edges' distances $dist_{ij}$ using the Floyd–Warshall algorithm [32–34].

Finally, the parameter $diam$ represents the maximum district diameter, that is, the maximum allowed distance separating two areas belonging to the same district. Therefore, without loss of generality, all edges (i, j) such that $dist_{ij} > diam$ are removed from the set E.

2.2. Partitioning variables and constraints

In the problem considered, the main decisions are represented in the model by the following variables:

- $assign_{ik} = \begin{cases} 1 & \text{if area } i \text{ is assigned to district } k \\ 0 & \text{otherwise} \end{cases}$
- $time_{ik} \geq 0$, time allocated to area i by district k.

Variables $assign_{ik}$ define the assignment of areas to districts, while variables $time_{ik}$ represent the district patrolling time allocation to each area.

The decisions concerning the definition of the districts and the allocation of the service time to the areas are modeled through the following constraints.

$$\sum_{k=1}^{K} assign_{ik} = 1 \qquad \forall i \in N \tag{1}$$

$$time_{ik} \leq cap \cdot assign_{ik} \quad \forall i \in N, \ k = 1, \ldots, K \tag{2}$$

$$\sum_{i \in N} time_{ik} = cap \qquad \forall k = 1, \ldots, K \tag{3}$$

$$assign_{ik} \in \{0, 1\} \qquad \forall i \in N, \ k = 1, \ldots, K \tag{4}$$

$$time_{ik} \geq 0 \qquad \forall i \in N, \ k = 1, \ldots, K \tag{5}$$

Constraints (1) state that all areas must be assigned to exactly one district. Constraints (2) impose that an area can receive patrol time only from its assigned district. Constraints (3) enforce the allocation of the time capacity to areas in such a way that all the time capacity available in a district is distributed among its assigned areas. The last two sets of constraints present the existence conditions for variables $assign_{ik}$ and $time_{ik}$.

2.3. Contiguity and diameter

The districts are required to be geographical contiguous (i.e., in each district one patrol could go from one area to another without leaving the district) and compact (i.e., closely and neatly packed together). The former requirement is for practical reasons, while the latter is for efficiency. In fact, a compact district allows patrols to reach every point quickly in case of an emergency. Compactness is achieved by limiting the maximum diameter of the districts, where the diameter is the maximum shortest-path distance between every pair of areas belonging to the same district. On the other hand, contiguity is imposed via a set of flow constraints, as originally proposed by Shirabe [35,36]. Differently from Shirabe's original formulation, in the present model each district

has a different network flow model associated. The variables required are the following:

- $source_{ik} = \begin{cases} 1 & \text{if area } i \text{ is the source of the flow} \\ & \text{corresponding to district } k \\ 0 & \text{otherwise} \end{cases}$

- $flow_{ijk} \geq 0$, flow from area i to area j in district k.

The model assigns to each district a source area (represented by variables $source_{ik}$), from which a flow (modeled by variables $flow_{ijk}$) originates.

The limit on the districts' diameter, on the other hand, is imposed by forbidding that two areas having a shortest-path distance ($dist_{ij}^{SP}$) greater than the maximum diameter ($diam$) can belong to the same district. Let $\overline{E} \subseteq E$ be the set of edges $(i, j) : dist_{ij}^{SP} > diam$. The constraints corresponding to these requirements are presented in the following.

$$flow_{ijk} \leq (|N| - 1) \cdot assign_{ik} \quad \forall k = 1, \ldots, K, \ (i, j) \in E \tag{6}$$

$$flow_{ijk} \leq (|N| - 1) \cdot assign_{jk} \quad \forall k = 1, \ldots, K, \ (i, j) \in E \tag{7}$$

$$\sum_{i \in N} source_{ik} = 1 \quad \forall k = 1, \ldots, K \tag{8}$$

$$source_{ik} \leq assign_{ik} \quad \forall k = 1, \ldots, K, \ i \in N \tag{9}$$

$$|N| \cdot source_{ik} + \sum_{j:(j,i) \in E} flow_{jik} \geq$$
$$\sum_{j:(i,j) \in E} flow_{ijk} + assign_{ik} \quad \forall k = 1, \ldots, K, \ i \in N \tag{10}$$

$$assign_{ik} + assign_{jk} \leq 1 \quad \forall k = 1, \ldots, K, \ (i, j) \in \overline{E} \tag{11}$$

$$source_{ik} \in \{0, 1\} \quad \forall k = 1, \ldots, K, \ i \in N \tag{12}$$

$$flow_{ijk} \geq 0 \quad \forall k = 1, \ldots, K, \ (i, j) \in E \tag{13}$$

Constraints (6) and (7) specify that an edge is part of a district's network-flow model only if both its origin and destination areas belong to the district, respectively. Next, constraints (8) and (9) define the source of a district's flow: each district can have only one source area which has to belong to the district. Flow conservation is enforced by constraints (10). In general, for each area and district, all the in-flow (i.e., flow generated in the area plus flow entering the area from the edges in the district's network-flow model) must be greater than or equal to the out-flow (i.e. flow leaving the area through the edges in the district's network-flow model plus flow consumed in the area). In particular, the district's source area (i.e. $source_{ik} = 1$) generates $|N|$ flow unit (i.e., an upper bound to the possible size of the district; see the first term of the left-hand side of the inequality). Also, each area in the district (i.e., $assign_{ik} = 1$) consumes flow unit (see the second term of the right-hand side of the inequality). Therefore, to be satisfied, constraints (10) requires all the areas belonging to the district to be connected directly or indirectly to the source areas, using only edges belonging to the district's network-flow model (as specified by constraints (6) and (7)), so that they can receive the flow unit consumed. The diameter conditions are enforced by constraints (11) *that ensure two areas* cannot belong to the same district if their shortest-path distance is greater than the maximum diameter. Finally, (12) and (13) specify the existence conditions for variables $source_{ik}$ and $flow_{ijk}$, respectively.

2.4. Fairness criteria and objective

Given a district configuration, the crime risk and population distribution attributes can be used to calculate the patrol service goal for an area i, the police contact goal for a population group p, and the risk exposure goal for a district k:

$$goal_i = (K \cdot cap) \cdot \frac{risk_i}{\sum_{j \in N} risk_j} \quad \forall i \in N \tag{14}$$

$$goal_p = (K \cdot cap) \cdot \frac{\sum_{i \in N} pop_{ip}}{\sum_{i \in N} \sum_{q \in P} pop_{iq}} \quad \forall p \in P \tag{15}$$

$$goal_k = \frac{1}{K} \cdot \sum_{i \in N} goal_i \cdot risk_i \quad \forall k = 1, \ldots, K \tag{16}$$

The patrol service goal (14) proportionally distributes the total patrolling time capacity available (i.e., $K \cdot cap$) among the areas, according to their risk. Similarly, the police contact goal (15) proportionally assigns the total patrolling time capacity among the population groups according to their size. These goals ensure that riskier areas and larger population groups have higher goal values. Finally, the risk exposure goal (16) equally divides the total risk exposure (i.e., $\sum_{i \in N} risk_i \cdot goal_i$) among the districts. The rationale is that a fair district configuration should be homogeneous across the districts in terms of risk exposure.

Given a solution and its specific time allocation (i.e., variables $time_{ik}$), the levels corresponding to the goals above can be computed as:

$$level_i = \sum_{k=1}^{K} time_{ik} \quad \forall i \in N \tag{17}$$

$$level_p = \sum_{i \in N} \left(\frac{pop_{ip}}{\sum_{q \in P} pop_{iq}} \sum_{k=1}^{K} time_{ik} \right) \quad \forall p \in P \tag{18}$$

$$level_k = \sum_{i \in N} risk_i \cdot time_{ik} \quad \forall k = 1, \ldots, K \tag{19}$$

The patrol service level (17) is the total patrolling time received by an area across all districts. The police contact level (18) is the total patrolling time received by a group assuming that, within one area, the patrolling time is distributed among the population groups proportionally to their size (i.e., equally among the individuals). Finally, the risk exposure level of a district (19) is the sum of the risk-weighted patrolling time allocated to the patrolled areas.

The fairness criteria values are calculated from the levels above and their corresponding goals:

$$W^T \leq \frac{level_i}{goal_i} \quad \forall i \in N \tag{20}$$

$$W^R \leq \frac{level_p}{goal_p} \quad \forall p \in P \tag{21}$$

$$W^W \leq \frac{level_k}{goal_k} \quad \forall k = 1, \ldots, K \tag{22}$$

In particular, the territorial fairness criterion (W^T) is the minimum ratio across all the areas of their patrol service level and their goal (20); the racial fairness (W^R) is the minimum ratio across all the population groups of their police contact level and their goal (21); finally, the workload fairness (W^W) is the minimum ratio across all the districts of their risk exposure level and their goal (22). As all the criteria are dimensionless ratios based on a goal value, they are comparable.

The decision-maker can express their preferences on the fairness criteria by setting the value of the preference weights w^T, w^R, and w^W, such that $w^T, w^R, w^W \geq 0$ and $w^T + w^R + w^W = 1$. The total fairness value, W, is computed as the preference-weighted sum of the fairness criteria:

$$W = w^T \cdot W^T + w^R \cdot W^R + w^W \cdot W^W \tag{23}$$

2.5. Formulation

The complete formulation can now be provided.

$$\max \quad W$$

Fairness Criteria and Objective : (17)–(23)

Partitioning Constraints : (1)–(5)

Contiguity and Diameter Constraints : (6)–(13)

The model maximizes the total fairness value, subject to partitioning, contiguity, and diameter constraints. The Pareto frontier can be

obtained by solving the model for different configurations of the values of w^T, w^R, and w^W.

Let K be the number of districts to be defined, $|N|$ be the number of areas, $|P|$ be the number of population groups, and $|E|$ be the number of edges. Then, the model includes: $(2K|N|)$ binary variables, $(K|N|+K|E|+|N|+|P|+K+4)$ real variables and $(3|N|+3|N|K+4K+2|E|K+2|P|+K|\overline{E}|)$ constraints (excluding the existence conditions on the variables).

As previously stated, the model is built upon the one in Liberatore et al. [24]; however, it presents significant differences that are highlighted in the following. (i) It considers three fairness criteria (i.e., territorial, racial, and workload) instead of just two. (ii) Contiguity is imposed by a set of flow constraints instead of using a Minimum Spanning Tree model; this results is a smaller and easier to solve model. (iii) It does not differentiate between the travel time between areas and patrolling time, as the travel time is still patrolling time in terms of crime deterrence. (iv) A maximum district diameter is explicitly imposed. (v) It provides a clear definition for the criteria goals; this was left to the decision-maker in Liberatore et al. [24]. (vi) The model maximizes the linear combination of the minimum fairness criteria values. The objective function in Liberatore et al. [24] is based on the Compromise Programming paradigm and minimized the convex combination of the scaled distances of the fairness criteria to their ideal. The new objective function is easier to interpret and is more computationally efficient, as it does not require calculating the ideal and the anti-ideal values of the criteria.

3. Computational analysis

This section introduces the analysis that studies the impact of an instance's characteristics and the decision-maker preferences on the problem solvability and the trade-off between the fairness criteria. Therefore, the goal of the computational analysis is to tackle the following research questions:

RQ1 What are the factors that affect the solvability and the solution time of an instance?

RQ2 What are the factors that affect the trade-off between the fairness criteria?

The dataset used in the experiments is presented in Section 3.1. RQ1 and RQ2 are addressed in Section 3.3 and Section 3.4, respectively, where the analysis conducted and the insights obtained are discussed and presented in detail.

3.1. Dataset

The dataset used in the analysis is generated as described in Appendix A and is comprised of 30 instances. All the instances have the same graph topology (i.e., 100 nodes and 267 edges, represented in Figure 5, Supplementary Materials) and we only consider two populations: a majority and a minority one. This has been done for the sake of simplicity and to limit the number of confounding factors, as the analysis focuses on the fairness criteria and the following instance characteristics: minority population ratio, Gini index, and minority population/crime risk correlation. In fact, these variables provide a description of the population group distributions and their relationship to the crime distribution, which are the main characteristics affecting the racial fairness criterion. The population ratio expresses how large the minority population is in comparison to the majority one; therefore, this attribute determines the total minority population. The Gini index is an evenness measure of segregation that can be used to measure residential segregation between a majority and a minority population. It varies between 0.0 and 1.0, where 1.0 indicates maximum segregation. It has been chosen over other measures because it satisfies the

Table 2
Summary statistics for the computational results.

	primal	dual	gap	time (s)	solved
Min.	0.2166	0.2166	0.0001	0.276	–
1st Qu.	0.8674	0.8861	0.0011	0.9168	–
Median	0.9524	0.9733	0.0066	2.5575	–
Mean	0.9274	0.9377	0.0199	172.5539	0.4044
3rd Qu.	1.0000	1.0000	0.0236	21.7490	–
Max.	1.5348	1.5348	0.2156	2675.8470	–
St. Dev.	0.1941	0.1929	0.0322	460.8884	–

four criteria established by James and Taeuber [37]: organizational equivalence, size invariance, transfers, and composition invariance. Its formula is provided in Eq. (A.2), Appendix A.2. Finally, the correlation between the minority population and the crime risk distribution measures the strength of the relationship between the risk of crime and the absolute size of the minority population in an area.

Again, to reduce the number of confounding factors in the experiments carried out the number of districts is fixed to $K = 5$. The preference weights w^T, w^R, and w^W, can take the following values: $\{0, 0.25, 0.5, 0.75, 1\}$. Given that $w^T + w^R + w^W = 1$, these values result in 15 feasible combinations of preference weights. All in all, 450 different instances are solved.

3.2. Implementation and instance solution

The code that processes the data and solves the optimization model is programmed in Python 3 [38]. The all-pair shortest-path distance matrix, $dist_{ij}^{SP}$ is computed by the Floyd–Warshall algorithm provided in the SciPy library [39]. Finally, the mathematical programming model is implemented in OR-Tools [40] and solved using Gurobi 9.1.2 (64-bit Linux) [41]. All the experiments are run on a Dell Precision 5540, equipped with Intel® Core™ i9-9880H CPU @2.30 GHz × 16 and 16 GB RAM. The standard configuration of Gurobi is used, which applies multi-threading and a relative MIP gap tolerance of 0.0001. A CPU time limit of 3600s is set on all the optimization processes.

The dataset and all the source code are freely available at [42].

For the sake of space, the full computational results are reported in Table 1 (Supplementary Materials), while a summary is given in Table 2. The *gap* statistics include only instances not solved within the time limit, while the computational time statistics include only the instances solved to optimality.

3.3. RQ1: Solvability and solution time

RQ1 concerns the solvability and the solution time. In the following, for each of the above, descriptive statistics are provided and an appropriate regression model is fit to understand the impact of the following attributes thereon:

- Minority population ratio (*prop*).
- Gini index (gini).
- Correlation between the minority population and the crime risk distributions (corr).
- Territorial and racial fairness preference (w^T) and (w^R) respectively. Note that the workload fairness preference, (w^W) is omitted from the models as it can be expressed as the linear combination of (w^T) and (w^R).
- Descriptive statistics for the risk distribution, i.e., minimum (min_r), first quantile ($q1_r$), median (med_r), third quantile ($q3_r$), maximum (max_r), median absolute deviation (mad_r), standard deviation (sd_r) and variance (var_r).

For each regression model, a backward stepwise procedure with BIC criterion is used to select the best subset of variables. The final models are provided in Tables 2 and 3 (Supplementary Materials).

Table 3
Summary statistics for the trade-off strength τ for each pair of fairness criteria.

	W^T VS W^R	W^T VS W^W	W^R VS W^W
Min.	0.00000	0.00000	0.00000
1st Qu.	0.05022	0.00577	0.00000
Median	0.07943	0.02561	0.00209
Mean	0.07559	0.03463	0.01053
3rd Qu.	0.10992	0.05364	0.01139
Max.	0.15517	0.11187	0.08943
St. Dev.	0.04322	0.03196	0.01865

Table 4
Coefficients associated with the instances' attributes in the trade-off strength τ regression models.

	W^T VS W^R	W^T VS W^W	W^R VS W^W
prop	–	−0.7647	−0.8379
gini	–	1.4480	1.6911
corr	1.0067	–	–

A dash (–) indicates that the coefficient is not statistically significant.

3.3.1. Solvability analysis

Overall 40.44% of the instances are solved to optimality. A logistic regression model is fit to explain the probability of an instance being solved to optimality. The resulting model includes w^T, w^R, mad_r, and sd_r. Both the fairness preference weights have negative coefficients, indicating that they both impact negatively the probability of solving an instance to optimality. However, the territorial fairness coefficient is five times higher than the racial fairness'; as these weights are expressed on the same scale, this indicates that increasing the territorial fairness preference is what reduces the likelihood of solving a problem to optimality the most. On the other hand, both measures of risk spread have a positive coefficient. Therefore, it can be inferred that the instances with higher variability in terms of crime risk are easier to solve.

3.3.2. Solution time analysis

According to Table 2 the average solution time of the optimally solved instances is 172.55s and the standard deviation is 460.89. An exponential regression model[2] is fit to explain the solution time. The model includes w^T, min_r, med_r, and mad_r. All the variables have positive coefficients, except for med_r. Therefore, the higher the decision-maker preference for the territorial fairness, the minimum risk value, and the risk median absolute deviation, the longer it will take to solve an instance. On the other hand, a higher median risk corresponds to instances with lower computational time.

3.4. RQ2: Fairness criteria trade-off

RQ2 concerns the trade-off between the fairness criteria. In particular, the focus is on quantifying the type of interaction between each pair of criteria and on understanding the impact of an instance's characteristics thereon. To this end, an ad-hoc quantitative measure of bi-objective trade-off strength, τ, is developed (see Appendix B for its definition). A strong trade-off implies that the frontier is 'far' from the ideal solution and the preferences of the decision-maker have a great impact on the objectives' values in the final solution. On the other hand, a weak trade-off implies that the frontier is 'close' to the ideal solution and that the preference weights do not have a large impact on the objectives' values, as long as they are strictly positive. In fact, in this scenario, a preference of zero in a fairness criterion would result in the worst possible value thereof, while a positive one, albeit very small, would provide an ideal (or almost-ideal) value with no (or little) impact on the other criterion's value.

3.4.1. Descriptive analysis of the trade-off strength

For every instance and fairness criteria pair (i.e., W^T VS W^R, W^T VS W^W, and W^R VS W^W) a trade-off strength value is computed using the solutions obtained. For the full list of results please see Table 4 (Supplementary Materials), while summary statistics are provided in Table 3. According to them, the trade-off among the criteria is very weak, with values ranging from 0 to 0.16. More in detail, W^R VS

---

[2] Time data usually follows an exponential distribution.

W^W scores on average 0.01, indicating that the frontier is generally almost right-angled and that the trade-off between the criteria is almost non-existent. On the other hand, the trade-off between W^T and W^W is slightly more nuanced but still very weak, as the average τ is 0.03. Finally, the average τ for W^T and W^R is 0.08, indicating that their trade-off is even more nuanced, albeit still weak. From these observations, we can derive the following:

- The weight for the workload fairness criterion, w^W, should be a positive value close to zero so that the model can take it into account and avoid assigning an anti-ideal value. In fact, according to the above, disregarding this criterion leads to an anti-ideal value, while any other preference value leads to the ideal or an almost-ideal value, without any detriment to the other criteria.
- The preference of the decision-maker has a stronger effect on the values of the territorial and racial fairness criteria (W^T and W^R, respectively).
- Given the low value of τ in all the instances, the overall impact of the decision-maker is limited, as long as all the preference weights are strictly positive to avoid having their values drop to the anti-ideal.

3.4.2. Regression model on the trade-off strength

As τ is a ratio, a beta regression model has been fit on the trade-off strength value for each pair of fairness criteria. Also, as $\tau \in [0, 1]$, prior to fitting the model its values have been transformed as suggested by Smithson and Verkuilen [43]:

$$\tau' = \frac{\tau \cdot (N - 1) + s}{N} \tag{24}$$

where N is the number of values (i.e., the number of instances, that is, 30 in the current dataset) and s is a constant between 0 and 1; the value recommended by Smithson and Verkuilen, $s = 0.5$, has been used. The final regression models are provided in Tables 5–7 (Supplementary Materials).

By observing the coefficients of the significant independent variables (p-value < 0.05) it is possible to draw insights into the impact of an instance's characteristics on the trade-off strength. In fact, significant variables with positive coefficients imply that the higher their value, the more impactful the preferences of the decision-maker. On the other hand, a negative coefficient suggests that the preferences of the decision-maker have a smaller impact on the fairness values of the final solution. The coefficients are summarized in Table 4. The proportion of the minority population on the majority (*prop*) has a negative impact on the strength of the trade-offs involving the workload fairness criteria (W^W). This suggests that the greater the relative size of the minority population, the less significant the preferences of the decision-maker in terms of workload fairness. On the other hand, the Gini index (gini) has got the opposite effect. The higher its value (i.e., the higher the segregation between the majority and the minority population), the higher the impact of the decision-maker on the final solution in terms of workload. In summary, if a territory has a large and non-segregate minority population, then the effect of the preferences of the decision-maker on the workload is minimal. On the other hand, the correlation between the minority population and the crime risk (corr) is the only attribute that affects the trade-off between the territorial and the racial fairness criteria (WT VS W^R). As the associated coefficient is positive,

it can be deduced that a higher correlation gives more importance to the preferences of the decision-maker. This is expected, as a high correlation implies that high crime risk corresponds to minority population presence, and vice-versa. Therefore, assigning patrolling resources to high-crime areas implies a high police exposition for the minority population.

4. Case study: Madrid central district

The efficacy of the model is validated on the Madrid Central District dataset and compared to the solution currently in use by the SPN. This dataset was first introduced in [9]. However, the present version underwent some minor adjustments in the graph topology, which are highlighted below. For the sake of completeness, the rest of the parameters are summarized below.

4.1. Dataset

The set of areas N is comprised of 111 census districts (see Figure 1a, Supplementary Materials), which are represented by their centroids. Two areas i and j are connected by an edge (i, j) if the corresponding census districts share part of their perimeter. As a result, graph G is comprised of 111 areas and 314 edges, as illustrated in Figure 1b (Supplementary Materials).

Parameters $dist_{ij}$, $\forall (i, j) \in E$, is computed as the great-circle distance between the areas' centroids in meters, i.e., the length of the arc linking the two points on a sphere.

The dataset provides three crime risk distributions, associated with three specific shifts:

- SATT3: Saturday, 10/13/2012, night shift (10PM–8AM).
- SUNT1: Sunday, 10/14/2012, morning shift (8AM–3PM).
- MONT2: Monday, 10/15/2012, afternoon shift (3PM–10PM).

The heat maps corresponding to each shift are represented in Figure 2 (Supplementary Materials).

Parameter $risk_i$ is the number of reported thefts occurred in area i during the shifts above. The patrolling time capacity for each district, cap, is defined as the duration of the shift in minutes, i.e., 600 for SATT3 and 420 for SUNT1 and MONT2. Concerning the number of districts, the value considered is $K = 6$. In fact, the standard patrolling configuration for the Central District of Madrid adopted by the SNP (see Figure 3, Supplementary Materials) partitions the territory into six patrolling districts, according to the administrative wards, which are then assigned to multiple officers.

The population data is extracted from the 2011 Spanish Census [44]. The data provides the number of people living in each census district according to their region of birth: (a) Spain; (b) other EU country; (c) European non-EU country; (d) Africa; (e) Caribbean, Central and South America; (f) North America; (g) Asia; (h) Oceania. Therefore, in the case study, the population group set P is comprised of eight elements, one for each of the above regions. The distribution of each population group is graphically represented in Figure 4 (Supplementary Materials).

The maximum district diameter, $diam$, is set to:

$$diam = \frac{\max_{(i,j) \in E} \left\{ dist_{ij}^{SP} \right\}}{\sqrt{K}} \tag{25}$$

resulting in $diam = 1156.51$ meters. These values have been checked and approved by an SNP expert which determined that this diameter value allows the agents patrolling a district to promptly intervene in case of emergency.

Finally, for the preference weights w^T, w^R, and w^W, the following value have been considered: $\{0, 0.25, 0.5, 0.75, 1\}$. Given that $w^T + w^R + w^W = 1$, these values result in 15 feasible combinations of preference weights.

4.2. Results and discussion

Overall, 45 instances are solved and the results are illustrated in Tables 8–10 (Supplementary Materials), while Tables 11–13 (Supplementary Materials) report the solution values corresponding to the standard patrolling configuration adopted by the SNP (Figure 3, Supplementary Materials). As the standard patrolling configuration only provides the partition of the graph into patrolling districts but not the patrolling time assigned to each area, the solutions have been computed by fixing the values of the variables $assign_{ik}$ according to the configuration and then solving the model. All the instances were solved to optimality.

Fig. 1 illustrates the trade-offs between pairs of fairness criteria for the case study. In particular, for each shift (i.e., SATT3, SUNT1, and MONT2) and pair of fairness criteria, a line chart of the criteria values is represented (as a continuous line), where each point represents the fairness values of a specific instance. To emphasize the relation between the incumbent pair of fairness criteria, only the solutions where the sum of their corresponding preference weights is equal to one are represented (e.g., the W^T vs W^R charts include only the fairness values of the instances where $w^T + w^R = 1$). Each plot also displays the line chart corresponding to the standard patrolling configuration (as a dashed line).

By observing the plots, it is possible to identify patterns. First, all the trade-offs that involve the workload fairness criterion (W^W) are not strong, while the trade-off between W^T and W^R is nuanced and, therefore, stronger. This result matches the conclusions obtained in Section 3.4.

To compare the solutions obtained by the model to that used by the SNP it is necessary to introduce the concept of dominance. Graphically, a plot dominates another if the latter is subjacent to the former. When the plots represent solution values that must be maximized (as in this case), this implies that the dominating plot represents a solution set that is better than the dominated one. From the plots in Fig. 1, it can be seen that the model's solutions (continuous line) always dominate the standard configuration (dashed line), except for SATT3 and W^R vs W^W, where they achieve equivalent solutions. All in all, these results validate the usefulness of the model in obtaining solutions with higher levels of effectiveness and fairness in a real-world context.

5. Conclusions

This paper introduces the first police districting model that includes territorial, racial, and workload fairness criteria. The model is empirically tested on a novel instance dataset and is applied to a real-world dataset about the Central District of Madrid, Spain. A trade-off analysis is carried out to understand the extent to which the fairness criteria interact with each other. After analyzing the results of the experiments, the following insights can be deduced:

- 40.44% of the dataset instances solved to optimality. However, the average gap for the non-optimal solution is 0.0199, indicating that the solution values are close to the optimal.
- Problem instances with a high crime risk spread are more difficult to solve, require more solution time, and result in higher solution gaps compared to instances with lower crime risk variability.
- A high territorial fairness preference weight negatively impacts the solvability and solution time of an instance.
- There exists a trade-off between the territorial and the racial fairness criteria. However, the trade-off strength is low, meaning that almost-ideal solutions can be achieved as long as both criteria are accounted for in the model.
- Optimal/semi-optimal workload fairness can be achieved without any detriment in the others, as long as its preference weight is strictly positive. In fact, if the preference weight is zero, the workload fairness value drops to its anti-ideal. This illustrates
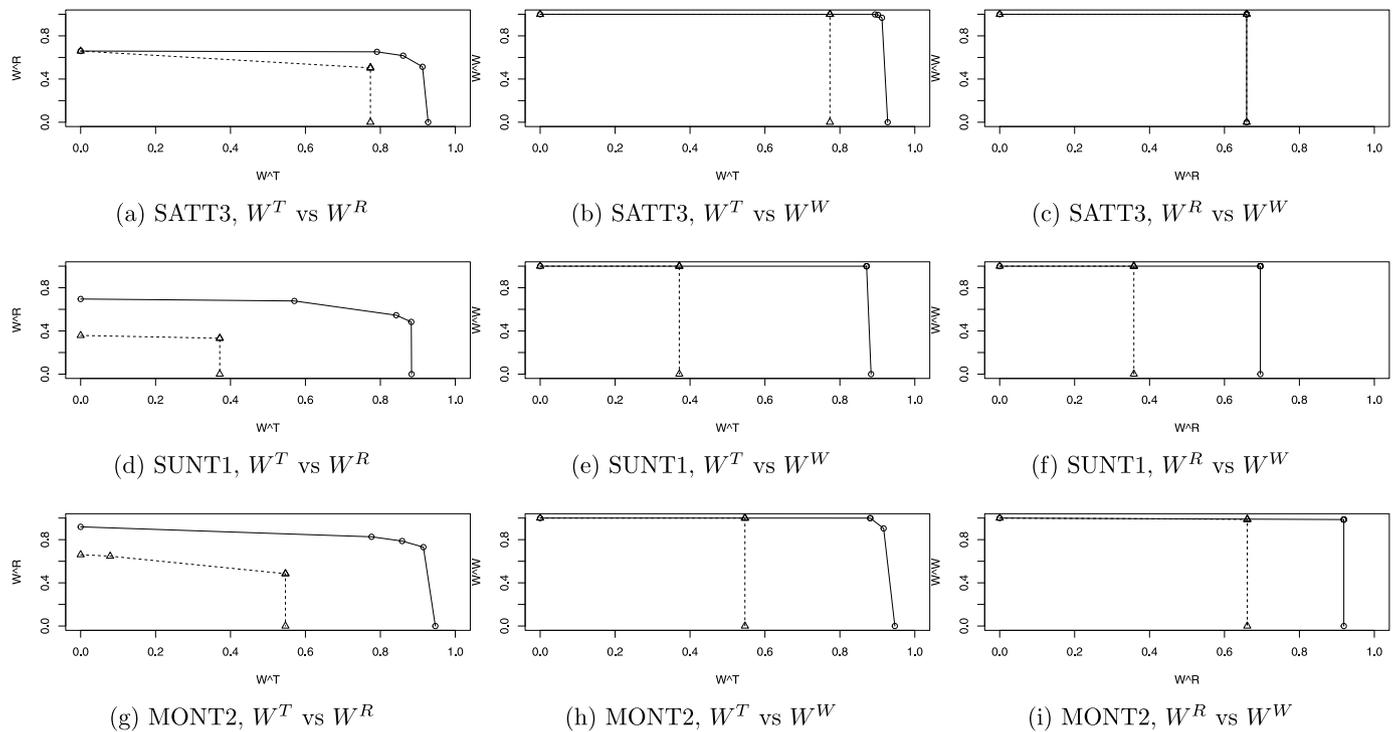
Fig. 1. Fairness criteria trade-offs representation. Model solutions are depicted with a circle and connected by continuous lines. The values of the standard patrolling configuration are represented with a triangle and connected by a dashed line.

the importance of including the workload criteria in patrolling models. It greatly benefits the agents without any detriment to the other stakeholders, while its omission would result in a significant workload imbalance across the districts.

• Therefore, it is recommendable that all the preference weights are strictly positive, with the workload fairness preference set to a value close to zero, as higher value would not have a major impact on the solution value.

• On the other hand, there is a strong trade-off between territorial and racial fairness. Also, this trade-off is stronger in instances where the minority population distribution is highly correlated to the crime risk distribution. In these scenarios, the preferences of the decision-maker really do have an impact on the fairness values of the patrolling configurations.

• A comparison between solutions obtained by the model and the districting configuration currently adopted by the SNP shows that the former outperforms the latter for all the criteria considered.

Despite the encouraging results, there is much that can be done to overcome the limits of the present study. The model proposed represents an abstraction of the patrolling operations that are carried out every day. For instance, it does not include all the operational aspects related to police districting and patrolling, such as the deployment of special police forces and the interactions with other public security bodies (e.g., Civil Guard and Local Police in Spain). Including these elements would make the model more realistic, applicable, and would benefit the analysis. Also, the model could be extended to account for the interests of a larger number of stakeholders (e.g., taxpayers and victims of crime). Finally, the model assumes that the distribution of the crime risk is deterministic and known *a priori*. However, when the risk measure is defined it is important to carefully consider all the potential biases involved. For example, police-reported crimes (e.g., originated from traffic stops) enforce conscious/unconscious biases rooted in the police agents (e.g., racial bias). On the other hand, reported crimes are affected by selection biases driven by unequal crime reporting rates across socioeconomic groups and geographical areas [45,46]. These biases could lead to further victimization, as

well as to over/under-reporting of specific areas, depending on their socioeconomic characteristics. In the model presented in this paper, bias effects are dampened by the presence of the racial fairness criterion, which favors solutions that maximize equity across population groups in terms of police exposition, regardless of the distribution of the crime risk. This limitation - which is shared among data-driven policing decision support tools and is an ongoing area of research [47,48] - could be overcome in a number of ways, including de-biasing the data used in the definition of the crime risk measure, or by considering uncertainty in the risk of crime. This would require formulating the model as a stochastic program. This is left as future research.

We hope that this work will inspire other researchers to extend the literature on social fairness in police operations and, more in general, in all the possible decision-making applications found in policing and in the public sector.

CRediT authorship contribution statement

Federico Liberatore: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Miguel Camacho-Collados:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Lara Quijano-Sánchez:** Data curation, Funding acquisition, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A link to the data/code is provided in the manuscript.

Acknowledgments

Appendix A. Dataset construction

This section explains in detail how the Police Districting Problem dataset has been built. In summary, the procedure randomly generates one graph topology based on Voronoi diagrams. Then, the population and risk distributions are defined according to three dimensions: minority population relative size, Gini index, and correlation between crime risk and minority population. To have a dataset that is representative of this three parameters space, 1000 distributions are generated, which are then clustered into 30 final instances that constitute the dataset. This procedure is detailed in the following. The final dataset and the code are freely available at [42].

A.1. Graph topology

Given a number of nodes N, the procedure used to build the graph is summarized in the following:

1. Randomly generate N points on a square bidimensional plane, using a uniform distribution U(0,100) for both coordinates. Each point corresponds to a node in the graph.
2. Use the point distribution to partition the square plane into cells according to its Voronoi diagram. Therefore, each node i has an associated cell and $area_i$.[3]
3. The edges are obtained from the Delaunay triangulation of the set of points, that is, the dual graph of the Voronoi diagram. The distance between each pair of nodes (i, j), $dist_{ij}$, is computed as the Euclidean distance between the corresponding points.

Figure 5 (Supplementary Materials) shows the topology of the graph used in the final dataset.

A.2. Population and crime risk distributions

All the dataset instances consider two populations ($|P| = 2$), a majority (population 1) and a minority one (population 2). Given a total majority population size, $POP_1 > 0$, and a proportion, $0 \leq \overline{prop} \leq 1$, the total minority population, POP_2, is equal to $(\overline{prop} \cdot POP_1)$. Therefore, the total population $POP = \sum_{p \in P} POP_p$ equals $\lceil (1 + \overline{prop}) \cdot POP_1 \rceil$. Also, the user can specify the desired values for the Gini index, $\overline{\text{Gini}}$, and the Pearson's correlation between the crime risk ($risk$) and the minority population distributions (pop_2), $\overline{\text{corr}}$. In particular, the Gini index varies between 0 and 1, representing no-segregation and completely segregated conditions, respectively.

It is possible to generate the population groups' and the crime risk's distribution data (i.e., pop_{ip} and $risk_i$, $\forall i \in N, p \in P$, see Section 2.1 and Table 1) by solving the following non-linear program.

$$\min_{pop, risk} Z(pop, risk) = \left| \overline{\text{Gini}} - \text{Gini}(pop) \right| + \left| \overline{\text{corr}} - \text{corr}(risk, pop_2) \right| \quad \text{(A.1)}$$

$$s.t. \ \text{Gini}(pop) = \frac{\sum_{i \in N} \sum_{j \in N} \left[pop_{i\cdot} pop_{j\cdot} \cdot \left| \frac{pop_{i2}}{pop_{i\cdot}} - \frac{pop_{j2}}{pop_{j\cdot}} \right| \right]}{2 POP^2 \frac{POP_2}{POP} \left(1 - \frac{POP_2}{POP} \right)} \quad \text{(A.2)}$$

---

[3] Although it is not used in this study, the area of the nodes is provided in the dataset for future uses.

$$\sum_{i \in N} pop_{ip} = POP_p \qquad \forall p \in P \quad \text{(A.3)}$$

$$\sum_{i \in N} risk_i = 1 \quad \text{(A.4)}$$

$$\sum_{p \in P} pop_{ip} \geq 1 \qquad \forall i \in N \quad \text{(A.5)}$$

$$risk_i \geq 10^{-4} \qquad \forall i \in N \quad \text{(A.6)}$$

$$pop_{ip} \geq 0 \qquad \forall i \in N, \ p \in P \quad \text{(A.7)}$$

Eq. (A.1) is the model's objective function, which minimizes the absolute distance of the Gini index and of the crime risk/minority population correlation to their respective desired values. Constraint (A.2) computes the Gini index [37] for any feasible value of pop. Note that $pop_{i\cdot} = \sum_{p \in P} pop_{ip}$. Next, constraints (A.3) establish that the sum of any population group across the areas must equal its given total. Similarly, constraint (A.4) imposes that the sum of the areas' risk must equal one. Therefore, the risk measure produced by the program can be interpreted as a probability. Constraints (A.5) enforce that the population in any area must be at least one. Finally, constraints (A.6) and (A.7) define the existence conditions for the variables $risk$ and pop, respectively, as well as their lower bounds.

The program above is non-linear and is solved using Sequential Least Squares Programming, which provides a locally optimal solution, initialized using randomly-generated values for pop and $risk$. Also, variable pop is continuous, which results in the population taking fractional values. These issues are addressed in the following subsection, where the complete instances generation procedure is presented.

A.3. Instances generation

An instance includes all the parameters that describe the topology of the graph, as well as the risk and the population distribution. The former is obtained as illustrated in Appendix A.1. On the other hand, the latter is generated according to Algorithm 1.

Algorithm 1 solveNLP procedure.

Input: $\overline{prop}, \overline{\text{Gini}}, \overline{\text{corr}}, POP, POP_1, POP_2$
Output: $pop^\star, risk^\star$
  $Z^\star \leftarrow \infty$
  $reps \leftarrow 10$
  while $reps \geq 0$ **do**
    $p\hat{o}p, \hat{risk} \leftarrow \text{argmin} \ \{(A.1)-(A.7)\}$
    $p\hat{o}p_1 \leftarrow \text{makeInteger}(p\hat{o}p_1, POP_1)$
    $p\hat{o}p_2 \leftarrow \text{makeInteger}(p\hat{o}p_2, POP_2)$
    if $Z(p\hat{o}p, \hat{risk}) < Z^\star$ **then**
      $Z^\star, pop^\star, risk^\star \leftarrow Z(p\hat{o}p, \hat{risk}), p\hat{o}p, \hat{risk}$
      if $|\overline{\text{Gini}} - \text{Gini}(p\hat{o}p)| < 5 \cdot 10^{-3} \land |\overline{\text{corr}} - \text{corr}(\hat{risk}, p\hat{o}p_2)| < 5 \cdot 10^{-3}$ **then**
        $reps \leftarrow 0$
      end if
    end if
    $reps \leftarrow reps - 1$
  end while

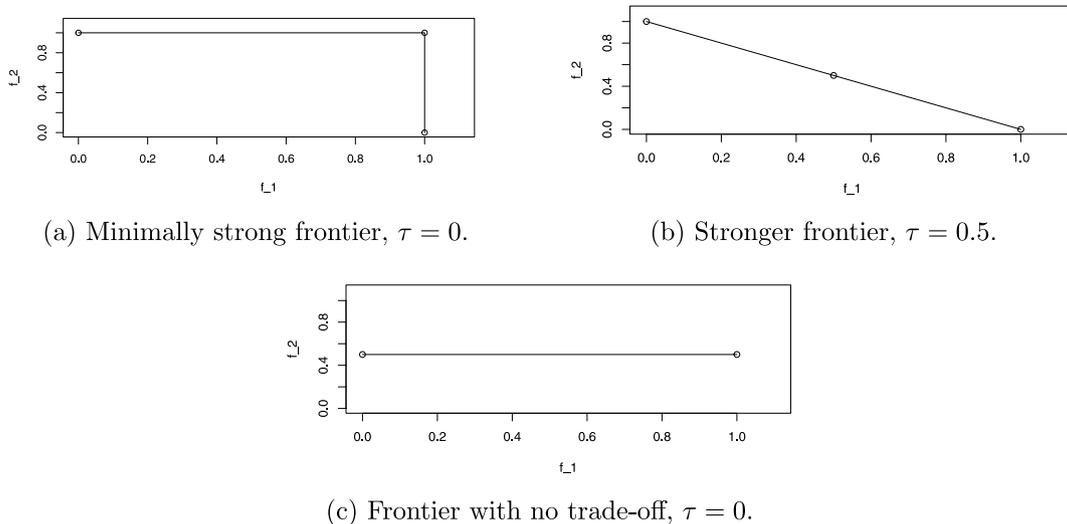**Algorithm 2** makeInteger procedure.

Input: vec, tot
Output: res
  $res \leftarrow \text{round}(vec)$
  $diff \leftarrow tot - \sum_{i \in N} vec$
  while $diff \neq 0$ **do**
    $ind \leftarrow \text{sample}(N, 1)$
    if $diff > 0$ **then**
      $res_{ind} \leftarrow res_{ind} + 1$
      $diff \leftarrow diff - 1$
    else if $diff < 0 \land res_{ind} > 0$ **then**
      $res_{ind} \leftarrow res_{ind} - 1$
      $diff \leftarrow diff + 1$
    end if
  end while

As previously mentioned, the formulation (A.1)–(A.7) is solved using Sequential Least Squares Programming, that is a local optimization

(a) Minimally strong frontier, $\tau = 0$.

(b) Stronger frontier, $\tau = 0.5$.

(c) Frontier with no trade-off, $\tau = 0$.

Fig. B.2. Sample approximate Pareto frontiers with different strength values. The axes represent the objective functions while the points are solutions on the frontier, drawn using lines. The values of the points and on the axes are by way of illustration.

algorithm. We initialize it by providing random feasible values for *pop* and *risk*. The goal Gini index and correlation values (i.e., $\overline{Gini}$ and $\overline{corr}$, respectively) might be impossible to achieve or hard to meet, therefore, solveNLP optimizes the formulation (A.1)–(A.7) at most 10 times and returns the best solution found. As the formulation is continuous, the solutions variables are converted to integers by the makeInteger procedure. makeInteger rounds a variable vector and increases or decreases by one the values of randomly chosen variables of the vector until the proper total is achieved (see Algorithm 2). Then, the Gini, correlation and solution values are updated. If the new solution value is lower than the best one found so far, then the best solution is updated. Also, if its Gini and correlation values are within a small threshold, $5 \cdot 10^{-3}$, of the desired values, then the search is interrupted. Finally, the best solution found is returned.

Algorithm 3 illustrates how multiple instances can be generated. The procedure core is the main loop, iterated *num_inst* times which, for the sake of the dataset generation, it is set to 1,000 to have a highly representative sample of all the possible population instances. In the loop, uniformly random values are assigned to $\overline{prop}$, $\overline{Gini}$, and $\overline{corr}$, while the majority population, POP_1, is a given parameter. Then POP_2 and POP can be computed. The procedure solveNLP (Algorithm 1) is then used to generate a problem instance satisfying as much as possible the requirements in terms of Gini index and minority/risk correlation. Without loss of generality, in the dataset it is assumed that the total majority population, POP_1, is 10,000.

Algorithm 3 Batch instance generation procedure.

Input: *num_inst*, POP_1
Output: D
    $count \leftarrow 0$
    $D \leftarrow \emptyset$
    while $count \leq num\_inst$ **do**
        $\overline{prop} \leftarrow U(0,1)$
        $\overline{Gini} \leftarrow U(0,1)$
        $\overline{corr} \leftarrow U(0,1)$
        $POP_2 \leftarrow \overline{prop} \cdot POP_1$
        $POP \leftarrow (1 + \overline{prop}) \cdot POP$
        $pop, risk \leftarrow$ solveNLP$(\overline{prop}, \overline{Gini}, \overline{corr}, POP, POP_1, POP_2)$
        $D \leftarrow D \cup (pop, risk)$
        $count \leftarrow count + 1$
    end while

A.4. Instances clustering and selection

The procedure presented in Algorithm 3 produces 10,000 instances. In this last step, the p most representative ones are chosen by clustering the instances into p groups, using the well-known p-median

problem [49]. The p-median problem chooses the p most representative instances (i.e., the medians) by minimizing the sum of the distance of each instance to its associated median. This method has been chosen to avoid the uncertainty introduced by other procedures, such as k-means, and because it clearly identifies a representative for each cluster.

The distance metric chosen is the Euclidean, where each instance is represented as a triplet of its main attributes, i.e., $(prop, Gini, corr)$. This procedure allows us to obtain a subset of instances that are evenly spread across the sample parameter space. The final dataset includes the $p = 30$ most representative instances. Table 14 (Supplementary Materials) illustrates the instances comprising the dataset and their corresponding attributes.

Appendix B. Quantitative Measure of Bi-Objective Trade-Off Strength

In the following, a quantitative measure of bi-objective trade-off is proposed. This measure provides a quantitative estimation of the strength of the trade-off between two objective functions. A trade-off is considered to be strong when the Pareto frontier is distant from the ideal solution. On the other hand, a trade-off is weak when it is close to the ideal solution.

It is important to note that the measure provided in this section is for the max–max case, that is, both objectives need to be maximized. The measure can be easily extended to the max–min and the min–min cases; however, a formal presentation of these cases is not provided as this is outside the scope of the present research.

Let $x \in X$ be the decision variables for a specific problem and let $f_1(x)$ and $f_2(x)$ be two objective functions that should be maximized. The Pareto frontier can be identified by optimizing the following joint objective function for all values of λ, with $0 \leq \lambda \leq 1$:

$$\max_{x \in X} \lambda \cdot f_1(x) + (1 - \lambda) \cdot f_2(x) \tag{B.1}$$

In practice, the frontier is usually approximated by solving the above for a representative subset of all the possible values of λ. Let us suppose that the Pareto frontier is approximated using L values of λ. This produces exactly L solutions, which correspond to L points in a $f_1 - f_2$ bi-dimensional space, i.e., $S_l \; \forall l = 1, \dots, L$. These can be plotted to obtain a graphical representation of the approximate Pareto frontier.

As can be seen in Fig. B.2, approximate Pareto frontiers usually present an 'elbow'. Intuitively, the closer it is to a right angle (Fig. B.2(a)), the weaker the trade-off between the objectives. In this case, it is possible to obtain the ideal solution (i.e., the solution having

maximum values in both objectives), while there exists an interaction between the objective functions. On the other hand, Fig. B.2(b) illustrates a strong trade-off, where the ideal solution (i.e., point $(1, 1)$ in the example) cannot be reached. Thus, the further the frontier is from the ideal solution, the stronger the trade-off is. Finally, Fig. B.2(c) shows an example where there is no trade-off between the two objectives. In this case, although the ideal point can be obtained, there is no interaction between f_1 and f_2. In fact, the value of f_2 is constant, irrespective to f_1.

Let points O, A, B, and C be defined as follows:

$$O = (\min_{x \in X} f_1(x), \min_{x \in X} f_2(x)) \tag{B.2}$$

$$A = (\min_{x \in X} f_1(x), \max_{x \in X} f_2(x)) \tag{B.3}$$

$$B = (\max_{x \in X} f_1(x), \max_{x \in X} f_2(x)) \tag{B.4}$$

$$C = (\max_{x \in X} f_1(x), \min_{x \in X} f_2(x)) \tag{B.5}$$

Points B and O are also known as the ideal and the anti-ideal point, respectively. Let *area* be the area of the polygon identified by the set of points $\{S_l \ \forall l \text{ in } 1, \ldots, L\} \cup O$. Following the intuition above, points A, B, and C define the strongest possible frontier for f_1 and f_2. Let *area** be the area of the polygon identified by the set of points $\{O, A, B, C\}$. Therefore, the measure of bi-objective trade-off strength, τ, can be computed as:

$$\tau = \begin{cases} 1 - \frac{area}{area^\star} & if \ area^\star > 0, \\ 0 & otherwise. \end{cases} \tag{B.6}$$

τ takes values between 0 and 1, where 1 indicates a maximally strong trade-off. Also, as it is expressed as a fraction, τ is dimensionless, which allows us to make consistent comparisons between trade-offs and frontiers resulting from different objective functions and solution spaces.

Appendix C. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.seps.2023.101556.

References

[1] Brantingham PJ, Valasik M, Mohler GO. Does predictive policing lead to biased arrests? Results from a randomized controlled trial. Stat Public Policy 2018;5(1):1–6.

[2] Saunders J, Hunt P, Hollywood JS. Predictions put into practice: A quasi-experimental evaluation of Chicago's predictive policing pilot. J Exp Criminol 2016;12(3):347–71.

[3] Johnson SD, Birks DJ, McLaughlin L, Bowers KJ, Pease K. Prospective crime mapping in operational context: Final report. London, UK: UCL, Jill Dando Institute of Crime Science; 2007.

[4] Perry WL. Predictive policing: the role of crime forecasting in law enforcement operations. Rand Corporation; 2013.

[5] Baycik NO, Sharkey TC, Rainwater CE. A Markov decision process approach for balancing intelligence and interdiction operations in city-level drug trafficking enforcement. Socio-Econ Plan Sci 2020;69:100700.

[6] Liberatore F, Camacho-Collados M, Vitoriano B. Police districting problem: Literature review and annotated bibliography. In: Optimal districting and territory design. Springer; 2020, p. 9–29.

[7] Camacho-Collados M, Liberatore F. A decision support system for predictive police patrolling. Decis Support Syst 2015;75:25–37.

[8] Camacho-Collados M, Liberatore F, Angulo JM. A multi-criteria police districting problem for the efficient and effective design of patrol sector. European J Oper Res 2015;246(2):674–84.

[9] Liberatore F, Camacho-Collados M. A comparison of local search methods for the multicriteria police districting problem on graph. Math Probl Eng 2016;2016.

[10] Ríos-Mercado RZ, editor. Optimal districting and territory design. International series in operations research & management science, vol. 284, Springer; 2020.

[11] Eiselt H, Marianov V. Maximizing political vote in multiple districts. Socio-Econ Plan Sci 2020;72:100896.

[12] Samanta S, Sen G, Ghosh SK. A literature review on police patrolling problems. Ann Oper Res 2021;1–44.

[13] Mohler G, Porter M, Carter J, LaFree G. Learning to rank spatio-temporal event hotspots. Crime Sci 2020;9(1):1–12.

[14] Sampson RJ, Wilson WJ. Toward a theory of race, crime, and urban inequality. In: Crime and inequality. Stanford, CA: Stanford University Press; 1995, p. 37–56.

[15] Braga AA, Weisburd DL. Does hot spots policing have meaningful impacts on crime? Findings from an alternative approach to estimating effect sizes from place-based program evaluations. J Quant Criminol 2020.

[16] Rinehart Kochel T. Constructing hot spots policing: Unexamined consequences for disadvantaged populations and for police legitimacy. Crim Justice Policy Rev 2011;22(3):350–74.

[17] Andrejevic M, Gates K. Big data surveillance: Introduction. Surveillance Soc 2014;12(2):185–96.

[18] Wikipedia contributors. George Floyd protests — Wikipedia, The free Encyclopedia. 2020, URL https://en.wikipedia.org/wiki/George_Floyd_protests. [Last Accessed 6 March 2023].

[19] Mohler GO, Short MB, Malinowski S, Johnson M, Tita GE, Bertozzi AL, et al. Randomized controlled field trials of predictive policing. J Amer Statist Assoc 2015;110(512):1399–411.

[20] Mohler G, Raje R, Carter J, Valasik M, Brantingham J. A penalized likelihood method for balancing accuracy and fairness in predictive policing. In: 2018 IEEE international conference on systems, man, and cybernetics. IEEE; 2018, p. 2454–9.

[21] Urcuqui C, Moreno J, Montenegro C, Riascos A, Dulce M. Accuracy and fairness in a conditional generative adversarial model of crime prediction. In: 2020 7th international conference on behavioural and social computing. IEEE; 2020, p. 1–6.

[22] Hart TC, Fitch CH. The predictive accuracy of prospective hot spot mapping and the race and ethnicity of street robbery victims: could a popular approach to crime fighting be a source of systemic racism? Crime Prevention Commun Saf 2022;1–16.

[23] Wheeler AP. Allocating police resources while limiting racial inequality. Justice Q 2019;1–27.

[24] Liberatore F, Camacho-Collados M, Quijano-Sánchez L. Equity in the police districting problem: Balancing territorial and racial fairness in patrolling operations. J Quanti Criminol 2021;1–25.

[25] Liao C, Scheuer B, Dai T, Tian Y. Optimizing the spatial assignment of schools to reduce both inequality of educational opportunity and potential opposition rate through introducing random mechanism into proximity-based system. Socio-Econ Plan Sci 2020;72:100893.

[26] Kistler A. Tucson police officers redraw division boundaries to balance their workload. Geography Public Saf 2009;1(4):3–5.

[27] Jingjing L, Fang Y, Tang N. A cluster-based optimization framework for vehicle routing problem with workload balance. Comput Ind Eng 2022;169:108221.

[28] Girit U, Azizoğlu M. Rebalancing the assembly lines with total squared workload and total replacement distance objectives. Int J Prod Res 2021;59(22):6702–20.

[29] Liu F, Niu B, Xing M, Wu L, Feng Y. Optimal cross-trained worker assignment for a hybrid seru production system to minimize makespan and workload imbalance. Comput Ind Eng 2021;160:107552.

[30] Zeng S, Wu Y, Yu Y. Multi-skilled worker assignment in seru production system for the trade-off between production efficiency and workload fairness. Kybernetes 2022.

[31] Chen H, Cheng T, Ye X. Designing efficient and balanced police patrol districts on an urban street network. Int J Geogr Inf Sci 2019;33(2):269–90.

[32] Floyd RW. Algorithm 97: Shortest path. Commun ACM 1962;5(6):345.

[33] Roy B. Transitivité et connexité. C R Acad Sci Hebd Seances Acad Sci D 1959;249(2):216–8.

[34] Warshall S. A theorem on boolean matrices. J ACM 1962;9(1):11–2.

[35] Shirabe T. A model of contiguity for spatial unit allocation. Geogr Anal 2005;37(1):2–16.

[36] Shirabe T. Districting modeling with exact contiguity constraints. Environ Plan B: Plann Des 2009;36(6):1053–66.

[37] James DR, Taeuber KE. Measures of segregation. Sociol Methodol 1985;15:1–32.

[38] Python Software Foundation. Python 3 documentation. 2008–, URL https://docs.python.org/3/.

[39] Jones E, Oliphant T, Peterson P, et al. SciPy: Open source scientific tools for Python. 2001, URL http://www.scipy.org/. [Last accessed 23 February 2023].

[40] Perron L, Furnon V. OR-tools. 2019, Google, URL https://developers.google.com/optimization/.

[41] Gurobi Optimization L. Gurobi optimizer reference manual. 2021, URL https://www.gurobi.com. [Last accessed 23 February 2023].

[42] Liberatore F, Camacho-Collados M, Quijano-Sánchez L. Dataset, source code and results for 'Towards Social Fairness in Smart Policing'. 2022, URL http://dx.doi.org/10.5281/zenodo.6583265. [Last accessed 23 February 2023].

[43] Smithson M, Verkuilen J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychol Methods 2006;11(1):54.

[44] INE, Instituto Nacional de Estadística (Spanish National Statistics Institute). Censos 2011, Resultados (census 2011, results). 2011, URL https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176992&menu=resultados&idp=1254735572981. [Last Accessed 23 February 2023].

[45] Xie M, Baumer EP. Neighborhood immigrant concentration and violent crime reporting to the police: A multilevel analysis of data from the national crime victimization survey. Criminology 2019;57(2):237–67.

[46] Buil-Gil D, Medina J, Shlomo N. Measuring the dark figure of crime in geographic areas: Small area estimation from the crime survey for England and Wales. Br J Criminol 2021;61(2):364–88.

[47] Alikhademi K, Drobina E, Prioleau D, Richardson B, Purves D, Gilbert JE. A review of predictive policing from the perspective of fairness. Artif Intell Law 2022;1–17.

[48] Purves D. Fairness in algorithmic policing. J Am Philos Assoc 2022;8(4):741–61.

[49] Hakimi SL. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. Oper Res 1965;13(3):462–75.

Federico Liberatore is a Lecturer at the School of Computer Science and Informatics, Cardiff University. He is a Fulbright Alumni and the recipient of several research awards. His main areas of expertise are Combinatorial Optimization and Data Science. His research focuses on mathematical models for policing, security, and personal safety. He is actively collaborating with law enforcement agencies on projects including predictive policing, smart patrolling, and femicides prevention. Also, Federico has relevant researcher experience in Disaster Management and Humanitarian Logistics. He has applied mathematical models to mitigate the effect of disasters and to plan for emergency response operations.

Miguel Camacho Collados is a Ph.D. in Mathematics and Statistics and a Fulbright Alumni. He joined the Spanish National Police in 2009 as an Inspector. In 2016, he took charge of the Statistics Unit. In 2017, he became the head of the Hate Crime National Office, Ministry of the Interior. Since April 2019, he coordinated the Artificial Intelligence Area at the Ministry of Economy. Currently, he works for the Spanish Council of State. Miguel has extensive experience both as a researcher and as a policymaker, as shown by his research publications and his participation in international organizations.

Lara Quijano-Sanchez, is a Ph.D. in Computer Engineering, an Associate Lecturer at Autónoma University of Madrid, and a member of the UC3M-Santander Big Data Institute of the Carlos III University of Madrid. She is specialized in data science, recommendation systems, user modeling, information search and retrieval, social network analysis, sentiment analysis techniques and design of predictive models using Big Data techniques. Lara is a well-known figure in the IA community where she has been able to develop a solid and significant research experience with several relevant contributions, published both in journals and international conferences.