

ABORDAGEM PARA OTIMIZAÇÃO DO PROCESSO DE ALOCAÇÃO DE VIATURAS POLICIAIS UTILIZANDO REGISTROS DE OCORRÊNCIAS CRIMINAIS PARA SUPORTE NO PROCESSO DECISÓRIO GERENCIAL NO ÂMBITO DA SEGURANÇA PÚBLICA.

Gláucio Bezerra Rocha

Centro de Informática - Universidade Federal da Paraíba
glauciobr@gmail.com

Gilberto Farias de Sousa Filho

Centro de Informática - Universidade Federal da Paraíba
gilberto@ci.ufpb.br

RESUMO

Em análise realizada no âmbito da segurança pública do Estado da Paraíba, um dos problemas identificados e trazido para esse trabalho foi a má distribuição de viaturas policiais, ocasionando atrasos e ineficiências quando em atendimento a população. Atualmente as viaturas policiais são distribuídas em bases fixas localizadas em bairros, onde praticamente cada bairro possui algumas viaturas a disposição para atender aquela área, e percebe-se que atual alocação dessas patrulhas não consideram algumas características do fenômeno, a exemplo de número de ocorrências existentes e distância de locais com maior incidência criminal. Esse trabalho propõe um processo de clusterização de áreas, para posteriormente prever a quantidade de crimes nessas regiões identificadas, e prever uma solução através da resolução do Problema de Localização de Máxima Cobertura Capacitada, propondo locais que viaturas policiais possam se situar para atender uma maior parte população, considerando menor distância, tempo e capacidade de atendimento, com valores mínimos aceitáveis.

PALAVRAS CHAVE. Clusterização, Series Temporais, Problema de Localização.

Tópicos (Problema de Localização, Clusterização)

ABSTRACT

In an analysis carried out within the scope of public security in the State of Paraíba, one of the problems identified and brought to this work was the poor distribution of police vehicles, causing delays and inefficiencies when serving the population. Currently, police vehicles are distributed in fixed bases located in neighborhoods, where practically each neighborhood has some vehicles available to serve that area, and it is clear that the current allocation of these patrols does not consider some characteristics of the phenomenon, such as the number of existing occurrences and distance from places with higher crime rates. This work proposes a process of clustering areas, to later predict the amount of crimes in these identified regions, and to prescribe a solution through the resolution of the Problem of Location of Maximum Capacity Coverage, proposing places that police vehicles can be located to serve a greater part of population, considering less distance, time and service capacity, with minimum acceptable values.

KEYWORDS. Clustering, Temporary Series, Location Problem.

Paper topics (Location Problem, Clustering)

1. Introdução

A segurança pública no Brasil, área de grande importância da sociedade atual, é um setor crítico nos Estados brasileiros. [Ferreira e Rigueira 2013] constatam que a segurança vive no limite de crises decorrentes de problemas estruturais, uma mídia pesada utilizada para denegrir sua imagem, e quanto mais a criminalidade aumenta os Estados não conseguem conter e controlar os problemas relacionados à segurança. É uma área de alta relevância pelo fato de estar diretamente ligada ao maior bem jurídico do ser humano, a vida, além do patrimônio. É responsável por proteger a vida, cessar e punir criminosos. De um lado temos um Estado desparelhado sem conseguir dar uma resposta a contento à sociedade, de outro, um sistema organizado controlado por criminosos, denominado de Crime Organizado. Em [Lum e Greene 2007] o crime organizado é definido como um empreendimento complexo, centralizado, constituído para a prática de atividades ilegais. Atualmente, para se combater o crime é necessário investimento em inteligência policial, por meio da Análise Criminal (AC). A análise criminal é responsável por produzir conhecimento policial, onde se avalia e acompanha sistematicamente o fenômeno da criminalidade em massa, instrumentando os operadores da segurança pública e proporcionando um suporte administrativo, tático, estratégico para a prevenção e repressão do crime e da violência [FERRO 2006].

No Estado da Paraíba a realidade não é diferente. Pouca evolução percebe-se na estruturação da segurança pública, principalmente no tocante a tecnologia da informação como ferramenta de gestão e solução para subsidiar a atividade de análise criminal na produção de inteligência policial. Observa-se a ausência de vários indicadores e/ou informações pertinentes e relevantes para a gestão de segurança pública que não estão acessíveis nos dias atuais para os gestores. Saber quais os crimes de maior incidência, bairros e/ou regiões mais perigosos, gêneros das vítimas, horários mais propícios do acontecimento de crime, locais que estão relacionados a determinados tipos de crimes, previsão de quantitativo de crimes em determinadas regiões para uma atuação mais planejada de efetivo policial, dentre outros questionamentos que não têm respostas, são exemplos de uma gestão precária, onde, caso o gestor da segurança pública tivesse posse desse conhecimento poderia gerir com maior eficácia um setor de tamanha relevância. A Polícia Civil, instituição ligada diretamente a segurança pública, e responsável por apurar todos os crimes ocorridos em território paraibano, no desenvolvimento de sua função de polícia judiciária, é responsável pelo registro, por meio das informações das vítimas, de todas as ocorrências de crimes ocorridas. Esses registros, que são formalizados em Boletins de Ocorrências (BO), contém todos os dados sobre o fato criminoso sofrido pela vítima, e caso bem utilizados e submetidos a uma metodologia de análise criminal pode responder muitas perguntas que hoje estão sem respostas, impactando diretamente na gestão e colhendo resultados mais assertivos no combate à criminalidade.

Em análise realizada no âmbito da segurança pública do Estado da Paraíba, um dos problemas identificados e trazidos para esse trabalho foi a má distribuição de viaturas policiais por toda a área geográfica das cidades, ocasionando atrasos e ineficiências quando em atendimento à população. Atualmente as viaturas policiais são distribuídas em bases fixas localizadas em bairros, onde praticamente cada bairro possui algumas viaturas a disposição para atender aquela área. Percebe-se que a atual alocação dessas patrulhas não considera algumas características do fenômeno criminal para serem definidas suas localizações, a exemplo de número de ocorrências existentes e distância de locais com maior incidência criminal. Dessa forma, é possível identificar bairros com maior índice criminal equipados com menos viaturas em detrimento de bairros com menor número de ocorrências, resultando em um superdimensionamento e/ou subdimensionamento de recursos e aparato policial, além de perceber que, as viaturas ficam circulando de forma aleatória nas suas respectivas regiões aguardando serem acionadas, tentando descobrir de forma empírica o local que ocorrerá um crime, e muitas vezes no momento do acionamento estão em locais adversos daqueles que normalmente ocorre crimes com maior frequência.

Considerando todo o cenário do problema supracitado, esse trabalho propõe uma abordagem que utiliza os milhares de registros de ocorrências policiais registrados pela Polícia

Civil que são ignorados, subutilizados, mais precisamente da cidade de João Pessoa-PB, que não são utilizados para fins de estudo do fenômeno do comportamento da criminalidade local, para fornecer um suporte a gestão da segurança pública. Para fins de gerar mais conhecimento policial, os dados históricos das ocorrências foram agrupados (clusterização), mapeando regiões com maior índice de criminalidade, independente de fronteiras legais existentes (bairros), mas com foco no número de ocorrências, apresentando a análise desses dados para que o gestor possa identificar padrões, tendência e sazonalidade por meio de séries temporais, e assim realizar previsões de ocorrências policiais nestas regiões identificadas no processo de clusterização, para um melhor planejamento. Posteriormente, identificadas as regiões com maiores índices de criminalidade e a previsão de crimes nessas regiões, propomos uma fase de prescrição da localização das viaturas policiais com o objetivo de cobrir as regiões com maior demanda criminal previsto pelas fases anteriores. Para isto, propomos um modelo matemático para o Problema de Localização de Máxima Cobertura Capacitada adaptado aos dados criminais, considerando distância de cobertura e capacidade de atendimento das viaturas.

2. Trabalhos Relacionados

[Silva 2014] apresenta em seu trabalho uma solução para posicionamento de viaturas policiais para um atendimento à população com maior eficiência e eficácia. O estudo, por meio das localizações geográficas de crimes ocorridos na cidade de João Pessoa, identificou áreas crônicas com maiores intensidades de crimes, através de uma maior concentração de ocorrências, e assim permitindo ao gestor uma melhor análise do problema e consequentemente a definição de estratégia tática e operacional mais adequada. Foram utilizadas técnicas como k-means e p-mediana para atingir os objetivos estabelecidos, assim como a implementação de duas meta-heurísticas, a GRASP combinado com VND e GRASP com um modelo exato, no qual obteve resultados muito semelhantes.

[Oliveira 2020] apresentou a importância de um rápido atendimento a população por parte do Serviço de Atendimento Móvel de Urgência (SAMU), em que o tempo de resposta desse serviço é relevante para o atendimento ser bem-sucedido, e para isso é importante que as viaturas do SAMU estejam bem localizadas ao ponto de atender a maior quantidade de usuários. O trabalho tem como objetivo fornecer um conjunto de ferramentas que possa possibilitar um melhor posicionamento de viaturas do SAMU aplicando o Problema de Localização de Máxima Cobertura.

[Ahmed et al. 2020] utilizam o Prophet, que é baseado em um modelo de aprendizado de máquinas, para prevê series temporais de energias de painéis fotovoltaico. A previsão confiável desses dados torna-se essencial para o planejamento da capacidade, em antecedência, para gerenciar com eficiência a distribuição de energia. Os dados gerados pelos painéis são submetidos ao modelo Prophet um dia antes para ser realizada a previsão de saída de energia. Concluiu-se que os dados coletados foram bastantes confiáveis.

[Darapaneni et al. 2021] realizou um estudo referente a uma campanha de vacinação, na Índia, onde tinha como objetivo prever o tempo mínimo necessário para vacinação da população para alcançar a imunidade de rebanho. Nesse estudo foi utilizado modelo SIR, que define a capacidade de disseminação da doença, e posteriormente, por meio de análise de séries temporais usando Prophet foi possível prever a quantidade de dias necessários para vacinar a população suficiente para atingir a imunidade de rebanho.

No estudo apresentado por [Jain e Prasad 2020] é relatado a importância do controle da medição de alguns parâmetros referente ao tráfego de redes utilizado na área de telecomunicação para prestar um serviço de melhor qualidade, permitindo que seja feito planejamento adequado os ativos de redes. O trabalho apresenta uma solução para esse cenário de previsão de parâmetros de performance de redes utilizando series temporais com o Prophet e o algoritmo XGBoost. É relatado pelos autores que o Prophet é um algoritmo capaz de realizar previsão utilizando grande quantidade de dados irregulares, ou anormal.

[Kumar e Pande 2021] destaca que técnicas de previsão de dados são utilizadas em vários âmbitos, a exemplo de vendas, bancos, saúde, mercado de ações etc, para resolver os diversos problemas relacionados, pois prever ajuda na tomada de decisão. O trabalho destaca algumas

ferramentas disponíveis com esse propósito, utilizando modelo de regressão e modelo exponencial logístico. É examinado alguns modelos de previsão, tais como o modelo aditivo, o modelo autorregressivo de média móvel integrada (ARIMA) e o modelo Prophet. O trabalho concluiu que o Prophet foi o modelo com melhor desempenho, pois apresentou um baixo índice de erro, melhor previsão e melhor ajuste.

3. Metodologia

A metodologia utilizada para o desenvolvimento desse trabalho é baseada em três pilares, que juntos compõe a solução proposta, quais sejam: Clusterização de dados de ocorrências policiais resultando na identificação de regiões e sua incidência criminal, previsão da quantidade de crimes por dia e semana nas regiões identificadas e bairros e a alocação inteligente de quantidade e localização de viaturas policiais para melhor atender as regiões/bairro e seus respectivos crimes.

3.1 Clusterização

Foram coletados mais de 300 mil registros de ocorrências policiais que ocorreram na cidade de João Pessoa-PB no intervalo de tempo de janeiro de 2017 a dezembro de 2021, ou seja, compreendendo um período de 5 (cinco) anos, por meio do Sistema de Procedimento Policial (SPP), que se trata de um software presente nas unidades policiais que registram os crimes sofridos pela população, materializado por meio de Boletim de Ocorrências (B.O). Cada registro de ocorrência representa um crime ocorrido. Nos dados coletados existem informações referentes ao crime perpetrado e da vítima, e principalmente do dia e local do delito. Por meio o local que o crime ocorreu, foi possível definir a latitude e longitude através do serviço de geolocalização da API do Googlemaps.

A clusterização é uma técnica de agrupamento de dados pertencentes a determinadas categorias que possuem um grande grau de similaridade entre si e maior dissimilaridade entre dados de outras categorias. Nesse trabalho utilizaremos o K-means para realizar a esta clusterização. Com objetivo de definir a melhor quantidade de clusters a serem formados, tendo em vista a importância desse valor para um melhor resultado do algoritmo K-means, considerando os dados de ocorrências criminais, latitude e longitude, utilizados nesse trabalho, esse estudo aplicará a técnica de Elbow Method [Humaira e Rasyidah 2021]. É importante ressaltar que [Shi et al. 2021] e [Kodinariya e Makwana 2013] destaca que o método Elbow tem um problema, enfatizando que por se tratar de uma técnica de análise unicamente visual, nem sempre permite distinguir de forma inequívoca, até por analistas experientes, o ponto com o menor ângulo formado, tendo em vista que em alguns casos a curva plotada ser bastante suave.

3.2 Series Temporais e Previsão de Crimes

De posse dos dados dos dias e regiões resultante do processo de clusterização, foram construídas suas respectivas séries temporais. As séries temporais foram formadas dispondo os dados por região, dia e quantidade de crimes. Com as séries temporais definidas, o próximo passo foi realizar a previsão da quantidade de crimes que poderiam ocorrer nas regiões identificadas por meio dessas séries. A previsão foi realizada por meio da *API Prophet*, que recebeu como entrada séries temporais de dados históricos por meio de dois campos, quais sejam: data e quantidade, representado pela linha azul do gráfico da Figura 1. Após a realização do treinamento, o *Prophet* apresenta uma previsão de uma série temporal de dias predefinidos e quantidades futuras, conforme ilustrado na linha vermelha do mesmo gráfico.



Figura 1. Previsão dos crimes em uma região baseado no seu histórico criminal.

3.3 Problema de Localização de Máxima Cobertura Capacitado

O modelo matemático discutido nessa seção visa fornecer a localização de patrulhas policiais em locais específicos considerando algumas restrições. Inicialmente se tem um conjunto de clientes I e uma quantidade de ocorrências a_i , além de um conjunto de facilidades K (patrulhas policiais) em possíveis locais J que possam atender essas ocorrências visando uma maior agilidade. Considerando s_k como a capacidade de atendimento de cada facilidade k e uma distância D como sendo o raio máximo pré-estabelecido que uma facilidade pode se encontrar de um ponto de demanda de clientes, busca-se o conjunto de locais onde essas facilidades devem ser postas, que é definido por N_i . O conjunto N_i é formado por todos os locais que $d_{ij} \leq D$, onde d_{ij} representa a função de distância do cliente i a localidade j , resultando em todas as localidades que se encontram dentro do limite máximo de distância D de cada demanda a_i .

Duas variáveis de decisão são empregadas no modelo, x_{jk} e y_{ik} . A variável x_{jk} , quando tiver seu valor igual a 1, indicará que uma facilidade k foi instalada em uma localidade j , caso contrário, seu valor será igual a 0. Já a variável y_{ik} representa a quantidade de demandas de um cliente i atendidas pela facilidade k .

Dessa forma a solução estabelece locais para posicionar as facilidades considerando a uma distância máxima pré-definida do cliente, assim como, limitar a quantidade de demanda a ser atendida por cada facilidade. Nesse contexto, apresentamos o seguinte modelo proposto e aplicado para suprir uma melhor otimização.

$$\text{Maximizar } \sum_{i \in I} \sum_{k \in K} y_{ik} \quad (1)$$

Sujeito a:

$$\sum_{i \in I} y_{ik} \leq s_k, \quad \forall k \in K, \quad (2)$$

$$\sum_{k \in K} y_{ik} \leq a_i, \quad \forall i \in I, \quad (3)$$

$$y_{ik} \leq a_i \cdot (1 - x_{jk}), \quad \forall i \in I, \forall k \in K, \forall j \notin N_i, \quad (4)$$

$$s_k \cdot \sum_{j \in N_i} x_{jk} \geq y_{ik}, \quad \forall i \in I, \forall k \in K, \quad (5)$$

$$\sum_{j \in J} x_{jk} \leq 1, \quad \forall k \in K. \quad (6)$$

A função objetivo (1) procura maximizar o número total de demandas (ocorrências policiais) atendidas por cada facilidade, por conseguinte, deve-se considerar as seguintes restrições.

A restrição (2) define que cada facilidade só pode atender s_k demandas dos clientes. A restrição (3) estabelece que cada cliente só pode ter a_i demandas atendidas pelas facilidades. A restrição (4) garante que os clientes só podem ser atendidos por facilidades atribuídas para a localidades cujo raio de cobertura máximo D seja atendida. A restrição (5) assevera que é obrigatório uma facilidade em uma localidade caso algum cliente seja atendido, como também que dada uma facilidade naquela localidade o número de demandas a_i que essa facilidade atende é menor ou igual a s_k . Por fim, a restrição (6) garante que cada facilidade será atribuída para no máximo uma localidade.

4. Resultados

As fases de clusterização e de previsão de crimes das regiões foram implementadas na linguagem Python, com o auxílio da heurística K-means, disponível na API Scikit Learn [Pedregosa et al. 2011], para as clusterizações e da API Prophet para realizar as previsões. Para implementar o modelo de alocação de viaturas proposto foi utilizada a linguagem de programação C++ com o compilador GNU GCC versão 9.4.0 (opção de compilação -O2) e auxílio do resolvidor matemático CPLEX na versão 20.1. Todos os experimentos foram executados em um computador equipado com processador AMD Ryzen 5700U com 8GB de RAM e executando o sistema operacional Ubuntu 20.04, limitados a 1 thread.

4.1 Clusterização

Para o processo de clusterização foi utilizado o algoritmo K-means, e quantidade de cluster K foi definido utilizando o método Elbow, mas como já discutido nesse trabalho o método Elbow tem seus desafios quando a curva formada é muito suave. Para superar a dificuldade de identificar o ponto ideal (cotovelo) que define o melhor K pelo método Elbow, tendo em vista ser um método exclusivamente visual, esse trabalho utilizou o método proposto por [Shi et al. 2021], que propõe um método quantitativo.

O primeiro passo do método apresentado por [Shi et al. 2021] é calcular a soma do erro quadrado (SSE, sigla em inglês) para cada valor de K . Com a lista dos valores de SSE para cada K , e considerando cada valor como um ponto, calcula-se o ângulo α formado pelas distâncias de três pontos adjacentes (P_i, P_j e P_k) através da equação $\arccos = \frac{a^2+b^2+c^2}{2ab}$, onde a é distância euclidiana entre os pontos P_i e P_j , b a distância euclidiana entre os pontos P_j e P_k e c sendo a distância euclidiana entre P_k e P_i . Todos os ângulos calculados são armazenados e no final é identificado o menor ângulo formado, sendo este o escolhido como melhor K para o conjunto de dados.

Foi submetido o conjunto de dados de ocorrências policiais e foi possível identificar o valor de K ideal igual a 3, como pode ser visto na Figura 2 que mostra os valores de SSE para cada valor de K . O Ponto de menor ângulo está marca com um X vermelho.

Considerando o melhor valor para $K = 3$, foi possível observar que o raio médio entre os pontos de cada cluster em relação ao seu centróide é de aproximadamente 9.7 km, sendo um raio de cobertura proibitivo para se atribuir a uma viatura. Sendo assim, foi estabelecido um raio médio de 4 km para cada cluster, sendo este valor atingido por clusters com pelo menos $K = 12$, como podemos observar na Figura 3.

Considerando a situação acima descrita, o conjunto de dados de ocorrências policiais foram submetidos novamente ao método Elbow, porém, de forma que considerasse a distância média entre os pontos e seus respectivos centroides de no máximo 4.0km de média, resultando no gráfico da Figura 4. Com o raio médio de 4.0km, o melhor ângulo formado pelo método Elbow indica um $K = 13$, ou seja, que o ideal para o cenário proposto é a formação de 13 clusters.

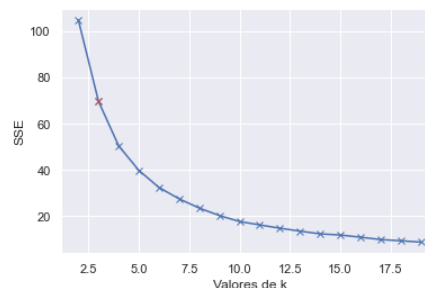


Figura 2. Gráfico Elbow indicando o melhor K igual a 3.

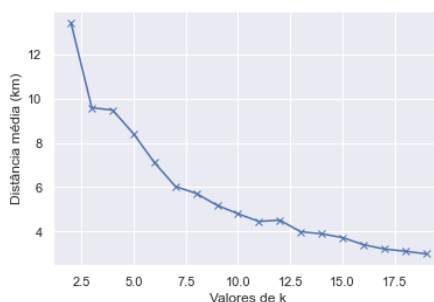


Figura 3. Relação do raio médio entre os pontos e os centroides dos clusters.

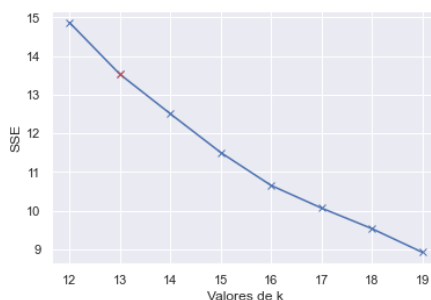


Figura 4. Gráfico de Elbow considerando o raio médio máximo de 4.0Km.

Os clusters formados podem ser observados na Figura 5.

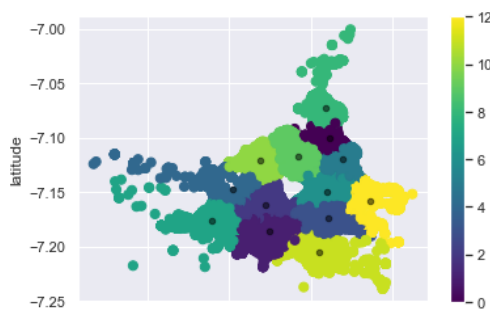


Figura 5. Clusterização dos dados considerando o número de cluster $K = 13$.

4.2 Series Temporais e Previsão de Crimes

Na presente seção serão comparados os resultados obtidos nas previsões da *API Prophet* sobre as séries temporais agrupadas por regiões, produto da fase de clusterização anterior e as séries temporais agrupadas por bairro. Para os fins de medição de desempenho do processo de previsão desse trabalho foi utilizado a métrica MAPE (Erro Percentual Médio Absoluto). O MAPE é uma métrica de simples entendimento, definida pela diferença média percentual entre os valores previsto e real de uma série temporal, ou seja, se o valor do MAPE foi 10% significa um erro de 10% no processo de previsão do valor previsto em relação ao valor real.

Na Tabela 1 são apresentados os valores de MAPE calculados de todas as regiões (coluna MAPE), já a coluna quantidade é referente a quantidade de crimes reais que ocorreram nas respectivas regiões contida em sua base de teste por fim a coluna erro absoluto traz a multiplicação do valor de MAPE por quantidade.

Tabela 1. MAPE ponderado das 13 regiões criminais.

Região	MAPE	quantidade	erro absoluto
1	38%	1483	563,54

2	18%	2607	469,26
3	28%	1665	466,22
4	20%	2706	541,20
5	14%	1467	205,38
6	25%	1636	409,00
7	19%	1602	304,38
8	14%	1006	140,84
9	35%	1125	393,75
10	25%	2243	560,75
11	20%	3005	601,00
12	12%	1490	178,80
13	31%	555	172,05
Total		22590	5006,15

O MAPE geral das regiões obtido com dados mensais foi de 22%, sendo computado pela razão entre o erro absoluto total de todas as regiões pelo total de crimes que realmente ocorreram ($\frac{5006,15}{22590} * 100 = 22\%$).

O mesmo procedimento foi aplicado aos dados agrupados por bairros, com todos os crimes e separando os crimes nos tipos de crimes contra o Patrimônio e crimes contra a Pessoa. Os resultados comparativos estão contidos na Tabela 2.

Tabela 2. Comparativo do MAPE em cada tipo de agrupamento dos dados.

Tipo de Crime	Agrupamento	MAPE
Todos	região	22%
	bairro	36%
Patrimônio	região	25,90%
	bairro	39,67%
Pessoa	região	36,69%
	bairro	68,29%

Em todos os cenários apresentados, considerando todos os crimes, ou separando por tipo, é possível identificar que ao utilizamos o agrupamento por região, o erro de previsão é no mínimo 14% menor comparado ao agrupamento por bairro.

4.3 Alocação de Viaturas utilizando o modelo matemático

Nessa seção serão apresentados os dados oriundos do processamento do modelo matemático proposto nesse trabalho. Para fins de processamento do modelo, alguns dados de entrada foram necessários, tais como: matriz de distância entre os centroides (latitude e longitude) dos bairros e regiões, quantidade de crimes previstos em cada região e bairro nos sete primeiros dias da semana do mês de janeiro de 2022, quantidade de viaturas disponível para patrulhamento em cada dia, capacidade de atendimento individual de cada viatura que fixamos em 5 crimes diários ($s_k = 5$) e por fim, o raio de cobertura das viaturas fixado em 4 km ($D = 4$).

As Tabelas 3 e 4 são formadas das seguintes colunas: coluna dia representa o dia do mês de janeiro de 2022 que foi realizada a projeção. As colunas $\sum a_b$ e $\sum a_r$ representam o somatório da quantidade de crimes previstos nos bairros e regiões, respectivamente. A coluna $|K|$ representa o número de viaturas que podem ser alocadas na solução. A coluna Z apresenta a quantidade de crimes coberto pela solução, coluna $T(s)$ representa o tempo computacional em segundos e $C(\%) = \frac{Z}{\sum a} * 100$ como o percentual de cobertura de crimes. Por fim, o $\Delta(\%)$ representa a diferença percentual de cobertura entre os valores obtidos da cobertura de regiões e bairros.

Tabela 3. Resultado de Alocação de Viaturas por dia.

dia	Instância			Bairros			Regiões			
	$\sum a_b$	$\sum a_r$	$ K $	Z	T (s)	C (%)	Z	T (s)	C (%)	Δ (%)
1	51	56	10	26	0,2	51	43	0	76,8	25,8
			11	28	0,2	54,9	46	0,1	82,1	27,2
			12	29	0,2	56,9	49	0,1	87,5	30,6
			13	30	0,2	58,8	52	0	92,9	34,1
			14	31	0,2	60,8	54	0,1	96,4	35,6
2	44	47	10	24	0,1	54,6	39	0,1	83	28,4
			11	25	0,1	56,8	41	0,1	87,2	30,4
			12	26	0,2	59,1	43	0,1	91,5	32,4
			13	27	0,2	61,4	45	0,1	95,8	34,4
			14	28	0,1	63,6	46	0,1	97,9	34,3
3	67	70	10	30	0,2	44,8	48	0,1	68,6	23,8
			11	32	0,2	47,8	52	0,1	74,3	26,5
			12	34	0,2	50,8	56	0	80	29,2
			13	36	0,3	53,7	60	0,1	85,7	32
			14	38	0,2	56,7	63	0,1	90	33,3
4	63	68	10	30	0,2	47,6	48	0,1	70,6	23
			11	32	0,2	50,8	51	0,1	75	24,2
			12	34	0,1	54	54	0,1	79,4	25,4
			13	36	0,2	57,1	58	0,1	85,3	28,2
			14	38	0,4	60,3	61	0,1	89,7	29,4
5	63	70	10	31	0,2	49,2	48	0,1	68,6	19,4
			11	33	0,2	52,4	52	0,1	74,3	21,9
			12	35	0,2	55,6	56	0,1	80	24,4
			13	37	0,2	58,7	60	0,1	85,7	27
			14	39	0,3	61,9	63	0,1	90	28,1
6	62	69	10	31	0,2	50	47	0,1	68,1	18,1
			11	33	0,2	53,2	51	0,1	73,9	20,7
			12	35	0,2	56,5	56	0,1	81,2	24,7
			13	36	0,2	58,1	59	0,1	85,5	27,4
			14	37	0,2	59,7	62	0,1	89,9	30,2
7	65	69	10	32	0,2	49,2	47	0,1	68,1	18,9
			11	34	0,2	52,3	51	0,1	73,9	21,6
			12	36	0,2	55,4	56	0	81,2	25,8
			13	38	0,2	58,5	59	0,1	85,5	27
			14	39	0,2	60	62	0,1	89,9	29,9

Na Tabela 3, a instância do dia 01 possui um total previsto para os bairros de 51 crimes e para as regiões de 56 crimes. São cinco soluções para este dia variando o número de viaturas $|K|$. Utilizando 10 viaturas no cenário de bairros, o modelo consegue apresentar um resultado que cobre 51% do total de crimes (26 dos 51 previstos), e gastou 0,2 segundos para fornecer essa solução, porém utilizando o mesmo número de viaturas, quando no cenário de regiões, o modelo apresenta uma solução que cobre 76,8% do total de crimes (43 dos 56 previstos), totalizando um aumento de 25,8% de cobertura dos crimes. Com o aumento de números de viaturas a serem utilizadas na solução, observa-se que há um aumento na cobertura de crimes quando utilizado o agrupamento de regiões obtendo até 96,4% da cobertura criminal.

Tabela 4. Resultado de Alocação de Viaturas por Semana.

Instância		Bairros			Regiões			
em	$ K $	(s)	(%)	(s)	(%)	(%)		
62	0	69	,6	6,7	81	,1	7,4	0,7
	1	79	,5	9,5	00	,1	2,7	3,2
	2	89	,6	2,2	19	,1	7,9	5,7
	3	98	,9	4,7	37	,1	2,8	8,1
	4	06	,1	6,9	49	,1	6,1	9,2
81	0	80	,5	7,2	92	,1	6,4	9,2
	1	92	,6	0,4	12	,1	1,7	1,3
	2	02	,7	3	32	,1	6,9	3,9
	3	12	,9	5,6	51	,1	1,9	6,3
	4	20		7,7	65	,1	5,6	7,9
17	0	01	,5	8,2	14	,1	4,9	6,7
	1	12	,7	0,8	37	,1	0,4	9,6
	2	23	,8	3,5	59	,1	5,7	2,2
	3	32		5,6	79	,1	0,5	4,9
	4	41		7,8	98	,1	5	7,2
43	0	09	,2	7,2	22	,1	2,5	5,3
	1	21	,8	9,9	48	,1	8,4	8,5
	2	31	,2	2,2	70	,1	3,3	1,1
	3	41	,7	4,4	91	,1	8,1	3,7
	4	50	,6	6,4	11	,1	2,6	6,2
51	0	08	,2	6,1	24		1,5	5,4
	1	20	,3	8,8	49	,1	7	8,2

2	30	,9	1	71	,1	1,9	0,9
3	39		3	91	,1	6,3	3,3
4	48		5	09	,2	0,3	5,3

Já a Tabela 4 são apresentados os resultados do modelo considerando as demandas previstas por semanas, totalizando 5 semanas no mês de janeiro de 2022. Como os dados são agrupados semanalmente, o valor anterior $s_k = 5$ que considerava que o número máximo de atendimento de ocorrências por viatura é de cinco crimes por dia, foi multiplicado por 7, adotando assim um $s_k = 35$.

É possível perceber na Tabela 4 que a qualidade das soluções geradas para o agrupamento por regiões continua sendo muito superior a soluções dos bairros, sendo todos os valores de $\Delta (\%) > 25$ e alcançando um $\Delta (\%) = 39,7$ no dia 1 com $|K| = 14$.

4.3 Conclusão

O presente trabalho apresentou uma solução que otimiza o processo de alocação de viaturas policiais tendo em vista o número previsto de crimes em áreas, tais como bairros e regiões, para uma melhor gestão operacional e estratégica por parte das gestões das unidades de policiamento operacional. No processo de clusterização foi utilizado o algoritmo K-means, no qual foram submetidas as ocorrências policiais e suas respectivas localizações (latitude e longitude). O algoritmo K-means tem como entrada a quantidade de cluster a ser considerado no processo de clusterização. Por meio do método Elbow (Cotovelo) e considerando o raio de distância máximo de 4 quilômetros entre os pontos, foi possível estabelecer a quantidade de regiões (clusters) ideal para o conjunto de dados submetido, apresentando como resultados 13 regiões.

Dando sequência a construção da solução foi realizada a fase de previsão de crimes, de forma separada, onde foi realizada a previsão considerando os 103 bairros e a previsão das 13 regiões. Após o processo de previsão, foi calculado o MAPE das regiões e bairros, com objetivo de mensurar a eficiência da previsão dos dados. Foi possível observar que, a perspectiva proposta por essa pesquisa, onde dividiu a área geográfica em regiões por meio da clusterização, é mais eficiente em relação a respectiva da divisão por bairros. Em todos os cenários apresentados obtendo um ganho mínimo de 14% na previsão dos crimes.

O último pilar da solução tratou do processo de alocação de viaturas policiais, considerando algumas restrições. Como resultado dessa fase é notório a diferença de performance quando a solução considera a perspectiva de região em detrimento de bairro, chegando a ter um ganho de cobertura de 35,6% a mais quando utilizado as regiões na previsão diária, e de 39,2% a mais quando na previsão semanal. É possível observar que enquanto o agrupamento em regiões tem soluções que apresentam mais de 90% de cobertura dos crimes previstos, o agrupamento por bairros possui soluções cuja cobertura não passam de 63.6% dos crimes previstos.

Por fim, é possível concluir que o presente trabalho apresentou como resultados valores bastantes satisfatório e de grande contribuição para ciência e pesquisa, já que em todos os cenários analisados é possível observar que a perspectiva de divisão da área geográfica por região apresentou dados mais performático em relação aos dados resultantes da divisão por bairro, possibilitando que sejam utilizados como direcionamentos para futuros trabalhos, ou até mesmo, para evolução deste.

Referências

Ahmed et al. (2020), "Forecasting PV Panel Output Using Prophet Time Series Machine Learning Model," 2020 IEEE REGION 10 CONFERENCE (TENCON) Osaka, Japan, November 16-19, 2020.

Darapaneni et al. (2021), "Forecasting Vaccination Drive In India for Herd Immunity using SIR and Prophet Model" 2021 IEEE World AI IoT Congress (AIIoT), 2021.

Ferreira, A. A. B. e Rigueira, D. L. A. (2013). Os indicadores-chave de desempenho como aliados da análise criminal. Revista Brasileira de Segurança Pública, [S. l.], v. 7, n. 2, 2013.

Ferro, A. L. (2006). Inteligência de segurança pública e análise criminal. Revista Brasileira de Inteligência. Brasília: Abin, v. 2, n. 2, abr. 2006.

Humaira H. e Rasyidah R., "Determining The Appropriate Cluster Number Using Elbow Method for K-Means Algorithm," ResearchGate, 2021, doi: DOI: 10.4108/eai.24-1-2018.2292388.

Jain, G. e Prasad, R. R. (2020), "Machine learning, Prophet and XGBoost algorithm: Analysis of Traffic Forecasting in Telecom Networks with time series data" 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Amity University, Noida, India. June 4-5, 2020.

Kodinariya T. M. e Makwana P. R, "Review on determining number of Cluster in K-Means Clustering," International Journal of Advance Research in Computer Science and Management Studies, 2013.

Kumar J. B. e Pande S. (2021), "Time Series Forecasting Model for Supermarket Sales using FB-Prophet," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 547-554, doi: 10.1109/ICCMC51019.2021.9418033.

Lum, C. M. e Greene, J. R. (2007). The Encyclopedia of Police Science. 3. ed. New York: Routledge, 2007. p. 283

Oliveira, C. P. de. (2020). Modelos de Otimização Aplicados ao Problema de Máxima Cobertura: Estudo de Caso do Samu-BH. Dissertação de Mestrado. Centro Federal de Educação Tecnológica de Minas Gerais, 2020.

Pedregosa et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830.

Silva, V. F. (2014). Problema de alocação de viaturas policiais: estudo de caso na cidade de João Pessoa-PB. 2014. 128 f. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal da Paraíba, João Pessoa, 2014.

Shi et al. (2021)., "A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm," EURASIP Journal on Wireless Communications and Networking, 2021.