LII Simpósio Brasileiro de Pesquisa Operacional
João Pessoa - PB, 3 a 5 de novembro de 2020

ESTRATÉGIA COMPUTACIONAL APLICADA A ALOCAÇÃO DE VIATURAS DE POLÍCIA NA CIDADE DE NATAL-RN

Elizeu Sandro da Silva
Rayana Souza Rocha
José Eric Francisco da Costa
Programa de Pós-graduação em Ciência da Computação - PPgCC
Universidade Federal Rural do Semi-Árido, UFERSA – Mossoró, RN – Brasil
Universidade do Estado do Rio Grande do Norte, UERN – Mossoró, RN – Brasil
{elizeusandro3, rayana.azus, jefc21}@gmail.com

Francisco das Chagas Lima Júnior
Carlos Heitor Pereira Liberalino
Programa de Pós-graduação em Ciência da Computação - PPgCC
Universidade do Estado do Rio Grande do Norte, UERN - Mossoró, RN - Brasil
{fclimajr, heitorliberalino}@gmail.com

RESUMO

A segurança pública é um dos aspectos essenciais para a garantia da ordem e da sobrevivência da população. No Brasil, é dever do estado garantir segurança pública a todos, preservando a ordem pública e a incolumidade das pessoas e dos patrimônios. Atualmente, a criminalidade é um dos fatores de mais impacto no Brasil, tendo este, 17 das 50 cidades mais perigosas do mundo, listadas em 2017. A fim de mitigar a criminalidade, a polícia como órgão executor da ordem pública, traça estratégias para reduzir as ocorrências. No entanto, tais estratégias não têm sido eficientes ao ponto de reduzir as taxas criminais em cidades capitais, como Natal-RN, cidade brasileira que ocupa a 4ª colocação em homicídios no mundo. Nesse contexto, uma possível solução para esse problema, é a redistribuição geográfica das viaturas de policiamento ostensivo em cidades que apresentam tais características, possibilitando uma melhor distribuição de cobertura das regiões de maiores ocorrências, sem deixar as demais descobertas. Sendo assim, objetivou-se com este trabalho, fazer um comparativo das respostas de algoritmos para os problemas de *p-medianas* e *p-centros* para alocação das viaturas de policiamento ostensivo, na cidade de Natal-RN. Foram utilizados dados georreferenciados de ocorrências de homicídios do ano de 2017, para a identificação das áreas de maiores intensidade criminais. A partir disso, foram aplicados os métodos de *clusters* e programação inteira, a fim de averiguar a aplicabilidade prática dos resultados encontrados. Por meio dos resultados obtidos, verificou-se uma similaridade entre as respostas dos métodos utilizados. Com isso, pode ser proposto a redistribuição das viaturas policiais de forma a cobrirem as áreas de maiores impactos criminais, sem que a distância entre as viaturas apresente maior insegurança aos policiais e apoiando nas tomadas de decisões estratégicas policiais.

PALAVRAS CHAVE. Segurança Pública, Homicídios, *k-means*.

Tópicos (D&SP – PO em Defesa e Segurança Pública, L&T – Logística e Transportes)

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

ABSTRACT

Public safety is one of the essential aspects for ensuring the order and survival of the population. In Brazil, it is the state's duty to guarantee public safety to all, preserving public order and the safety of people and property. At the same time, crime is one of the most impacting factors in Brazil, with 17 of the 50 most dangerous cities in the world listed in 2017. In order to mitigate crime, police as a law enforcement agency devise strategies for reduce the occurrence. However, such strategies have not been effective to the extent of reducing the default rates in capital cities, such as Natal-RN, a Brazilian city that occupies the 4th place in homicides in the world. In this context, a possible solution to this problem is the geographic distribution of ostensive police vehicles in cities with such characteristics, allowing a better coverage distribution of the regions with the highest occurrences, without leaving the other discoveries. Thus, the objective of this work was to compare the algorithm responses to the problems of dep-medianasep-centers for the allocation of ostensible policing vehicles in the city of Natal-RN. Georeferenced data of homicide occurrences from the year 2017 were used to identify the areas with the highest criminal intensity. From this, the decluster and integer programming methods were applied, in order to ascertain the practical applicability of the results found. By means of the obtained results, a similarity was verified between the answers of the used methods. With this, it is possible to propose the redistribution of police vehicles in order to cover the areas with the greatest impact, without the distance between the vehicles presenting greater insecurity to the police and supporting in the strategic police decision making.

KEYWORDS. Public Security, Homicide, *k-means*.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

1. Introdução

A segurança pública é um dos aspectos essenciais para a garantia da ordem e da sobrevivência da população. No Brasil, é dever do Estado garantir segurança pública a todos, preservando a ordem pública e a incolumidade das pessoas e do patrimônio BRASIL [1988]. Nos dias atuais, a criminalidade é identificada como um dos maiores problemas do Brasil, que possui 17 das 50 cidades mais perigosas do mundo BBC [2018]. Sendo o país com maior representatividade na lista e tendo o número de homicídios alarmante em 2015, ultrapassando o índice de 28,9 mortes a cada 100 mil habitantes IPEA [2017].

Dentre as regiões brasileiras, a que demonstra situação mais preocupante é a nordeste, que contém 11 (onze) das 17 (dezessete) cidades mais perigosas do mundo BBC [2018]. Sendo Natal-RN, a cidade brasileira que lidera a lista, ocupando em 2017 a quarta posição mundial com a taxa de 102,56 mortes e oitava em 2018, obtendo a taxa de 74,67 por cada 100 mil habitantes IPEA [2017]; BBC [2018]; CCSPJP [2019].

Com o intuito de mitigar a criminalidade, a polícia como órgão executor de ações intensivas e ostensivas tem utilizado centros integrados de atendimentos, que são utilizados para emergências e captação de informações, para suprimir as ações criminosas de Freitas Neto et al. [2014]. As informações adquiridas pelos centros de atendimentos, juntamente com os formulários de ocorrência policial presenciais, são utilizados pelos departamentos de segurança pública para traçar estratégias de combate a criminalidade. No entanto, tais estratégias não têm sido eficientes ao ponto de reduzir a taxa de ocorrências de crimes em cidades capitais, como Natal-RN.

Uma possível solução para este problema, é a redistribuição geográfica das viaturas de policiamento ostensivo dentro da cidade, possibilitando uma melhor distribuição de cobertura das regiões de maiores ocorrências, sem deixar as demais descobertas. Posicionar bens ou serviços, para o atendimento de demandas, estabelecendo critérios de eficiência, como distância ou tempo de resposta para serviços públicos emergenciais se enquadra no problema computacional conhecido na literatura como localização de facilidades da Silva et al. [2013].

Neste contexto, o Problema das *p-medianas* é um dos problemas clássicos de localização, que procura identificar, dentro de uma região pré-estabelecida, os melhores locais para se instalar um serviço, denominado facilidade. O objetivo é localizar p vértices (denominados medianas) em uma rede contendo n vértices, de forma a obter a menor soma das distâncias de cada vértice até a mediana mais próxima. Uma vez localizadas as múltiplas medianas, são formados *clusters* (agrupamentos), sendo que cada mediana é alocada a certo número de vértices da demanda Bezerra et al. [2008]. Da mesma forma que pode ser encontrado um conjunto solução para o problema de localização por meio do uso de *p-medianas*, o mesmo pode ser transformado em um problema *k-means*, quando o cálculo das distâncias podem ser calculados de forma contínua com o uso da métrica euclidiana quadrada Tan et al. [2006].

Sendo assim, o modelo proposto tem por objetivo otimizar a locação de viaturas policiais. As restrições garantem que cada cliente seja atribuído a alguma localidade, garantindo que a demanda total dos clientes designados por localidade não exceda a capacidade da instalação.

Desse modo, neste artigo é proposto uma modelagem computacional para o problema de distribuições de viaturas da polícia militar, baseado no modelo de localização de *p-medianas* e o modelo *k-means*, com o objetivo comparar os resultados obtidos afim de

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

otimizar a estratégia policial na prevenção de ocorrências.

2. Fundamentação Teórica

2.1. Problema das *p-medianas*

Problemas de localização e suas aplicações constituem uma importante linha de pesquisa operacional, que visam determinar a configuração de custo mínimo de instalação de facilidades. Tais problemas buscam determinar pontos estratégicos onde serão sediadas facilidades para atender, da melhor maneira possível, um conjunto espacialmente distribuído de pontos de demanda Bezerra et al. [2008].

Esses problemas podem ser classificadas como problemas de cobertura e de localização de medianas Arakaki e Lorena [2006]. O objetivo do problema de facilidades de *p-mediana* é localizar p instalações dentro de um conjunto de n pontos, tal que $(n > p)$, de forma que a soma das distâncias entre cada ponto de demanda e sua mediana/instalação mais próxima seja minimizada Rabie et al. [2014]. Sendo considerado um problema complexo de natureza combinatória para os quais não se conhecem algoritmos polinomiais capazes de obter a solução exata, exigindo grande esforço computacional e tornando-os de difícil solução por meio de algoritmos exatos, em que, mesmo utilizando métodos heurísticos e de agregação especializados na solução desse problema faz-se necessário um considerável esforço computacional de Almeida [2014]; Rabie et al. [2014].

Sabendo disso, para trabalhar com *p-medianas*, os dados relevantes para o problema são: o número finito de pontos de demanda, o número finito de candidatos a instalação de facilidades, a distância entre os pontos de demanda e o número p de facilidades a serem instaladas Pereira [2005].

O modelo matemático do problema de localização de facilidades com limites para o número de instalações abertas é considerado como um problema de programação inteira binária. As seguintes notações são introduzidas para formular o modelo matemático Dantrakul et al. [2014].

$$minZ = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \tag{1a}$$

$$s.a : \sum_{j=1}^{n} x_{ij} = 1, \forall i \in V \tag{1b}$$

$$\sum_{j=1}^{n} x_{ij} = p \tag{1c}$$

$$x_{ij} \leq x_{jj}, \forall i, j \in V, i \neq j \tag{1d}$$

$$x_{ij} \in \{0, 1\}\, \forall i, j \in V \tag{1e}$$

Onde,

$[c_{ij}]$nxn = Matriz simétrica de distância com $[c_{ij}] = 0$, para todo i.

$[x_{ij}]$nxn = Matriz de alocação $x_{ij} = 1$, se toda a demanda do nó i é suprido a partir de uma instalação em j, e $x_{ij} = 0$ caso contrário.

$x_{ij} = p$ caso o ponto i seja uma mediana e $x_{ij} = 0$ contrário.

A restrição (1b) garante que todos os pontos do centroide i deverão ser atendidos por uma única viatura.

A restrição (1c) garante que o número de medianas a ser localizada será igual ao valor p. O número ótimo de *p-medianas* é um objetivo a ser alcançado no problema (*p-max* = 12).

A restrição (1e) corresponde às condições de integridade.

2.2. *K-means*

O *K-means* é um algoritmo de aprendizagem de máquina não supervisionado, utilizado para a predição de dados. Esse algoritmo cria um conjunto homogêneo de *clusters* que minimizam as distâncias entre os elementos de um *cluster* e seu centro Tsihrintzis et al. [2019], cada um desses grupos pode ser definido como uma região na qual a densidade de objetos é localmente mais alta do que em outras regiões Likas et al. [2003]. Depois de selecionar *k-centros* de *clusters* iniciais, cada ponto de dados é atribuído a um grupo com base em uma medida de distância calculada e, em seguida, os *k-centros* de *cluster* são recalculados repetidamente, até que seja encontrado um conjunto ideal de *k-clusters* com base em uma função objetiva Nasraoui e N'Cir [2018]. A função de custo a ser minimizada, baseia-se em Hair et al. [2009].

$$custo = \sum_{k=1}^{k} \sum_{x_i \in C_k} d(x_i, x_{0k}) \tag{2}$$

Onde x_{0k} é o centroide do *cluster* C_k e $d(x_i, x_{0k})$ é a distância entre os pontos x_i e x_{0k}. Desta forma, o algoritmo encontra um conjunto de *hosts* para atuar como centroides $C = \{c_1, c_2, ..., c_k\}$ aleatoriamente e a distância d, baseada no cálculo da distância Euclidiana, a métrica mais utilizada para análise de agrupamento Vicini e Souza [2005]. Sua formulação matemática é apresentada a seguir.

$$distância = \sum_{v=1}^{V} (y_i - c_k)^2 \tag{3}$$

Onde, $y_i - c_k$ é a distância ao quadrado entre y_i e c_k, e o centro mais próximo entre os *k-centros* obtidos até então (ou seja, o centro de *cluster* em que y_i e c_k pertencem) Likas et al. [2003].

3. Metodologia

Para a realização deste trabalho foram utilizados de dados a cerca dos homicídios registrados na cidade de Natal-RN. Tais dados, foram obtidos por meio de solicitação à Secretaria de Segurança Pública e da Defesa Social (SESED)[1], no Portal do Sistema Integrado de Informações ao Cidadão do Rio Grande do Norte (SICRN)[2]. A base de dados solicitada contem registros sobre os homicídios ocorridos na cidade de Natal-RN, no período de 2012 a 2017. Para o tratamento e análise dos dados, foram utilizadas ferramentas como a Linguagem R e o Ambiente de Desenvolvimento *RStudio*, às quais permitem a análise de dados e plotagem gráfica dos mesmos. Nos dados solicitados eram contidos registros referentes a localização dos homicídios, como logradouro e bairro, os quais no processo de transformação, foram exportados para um arquivo do formato ".csv", de forma a possibilitar

[1]Disponível em: <http://www.defesasocial.rn.gov.br/>
[2]Disponível em: <http://www.sic.rn.gov.br/>

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

a análise dos registros de interesse. Cada registro da tabela foi convertido para um ponto georreferenciado, com longitude e latitude. Esses pontos foram gerados no *RStudio* com o auxílio da biblioteca do *Google Maps*[3], a *ggmap*, obtendo então, uma matriz de dimensões [n:3], como apresentado na Figura 1. No total, 16 registros foram exclusos pela ggmap, por erro na realização do registro na base de dados.

Figure 1: Associação dos dados de localização com suas respectivas longitude e latitude

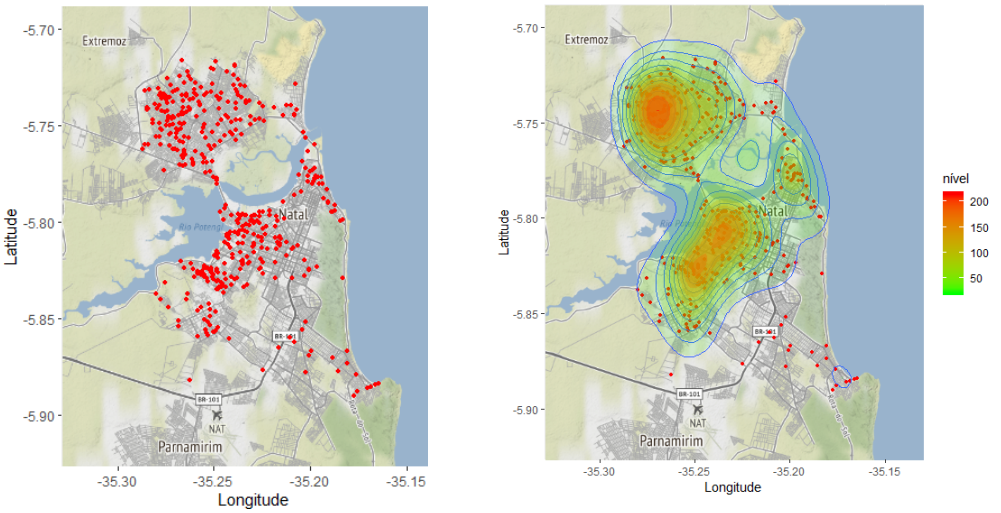
Filter

	LOGRADOURO.DA.OCORRENCIA	lon	lat
1	RUA NONATO MOTA COM ROCHA POMBO	-35.21581	-5.807899
2	AV. PRESIDENTE BANDEIRA, VILA SAO FRANCISCO EM FR...	-35.21743	-5.797088
3	RUA ARI PARREIRAS, Nº 1419, GALPAO DA TORCIDA ORG...	-35.22117	-5.794341

Fonte: Autoria própria (2019)

Com base no tratamento e a transformação dos dados realizadas na etapa anterior, foi gerado um mapa da cidade de Natal-RN, contendo os dados dos homicídios, onde foi adicionado um ponto no mapa para cada registro de localização, associados aos registros de homicídios do ano de 2017. A partir do gráfico obtido, é possível observar as regiões da cidade de Natal que obtiveram maiores índices de homicídios nesse período. Assim como visto na Figura 3(a), as áreas de maiores ocorrências de homicídios em 2017 foram as regiões Norte e Oeste. Essas informações podem ser observadas também na Figura 3(b), onde a partir do mapa de incidência criminal, foi gerado uma representação de mapa de calor para auxiliar na visualização das áreas de maior intensidade de homicídios no ano de 2017.

Figure 2: Análise gráfica dos dados de homicídios da cidade de Natal-RN no ano de 2017



(a) Mapa de ocorrências criminais da cidade de Natal-RN para o ano de 2017.

(b) Mapa de calor baseado nas ocorrências criminais da cidade de Natal-RN para o ano de 2017.

Fonte: Autoria própria (2019)

[3]Disponível em: <https://cloud.google.com/maps-platform?hl=pt-br>

Foi utilizado do *software CPLEX*, para a idealização de um modelo matemático inteiro para a resolução do problema de *p-medianas*, com 12 facilidades no máximo. O valor máximo foi obtido por meio de uma média, utilizando uma viatura para cobrir no máximo 3 bairros, ou seja, para os 36 bairros de Natal, seriam necessárias 12 viaturas. Esse valor é coerente com a realidade, segundo informações obtidas na Secretaria da Segurança Pública e da Defesa Social do RN.

Já para a aplicação do algoritmo *k-means* foi utilizado da ferramenta MATLAB[4]. Utilizando para a entrada do algoritmo o valor de k igual a 12. Os valores obtidos como respostas dos métodos, foram expressos em gráficos e são discutidos na Seção 4.

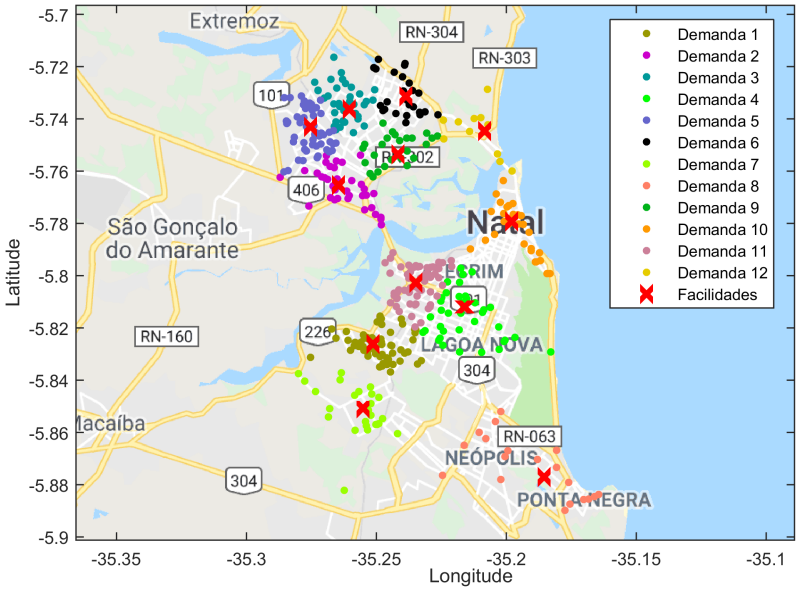
4. Resultados

Na Figura 3 é exibido a solução ótima gerada pelo modelo inteiro para o problema de *p-medianas*. Esta solução foi obtida pela implementação do modelo na ferramenta *CPLEX*. Cada conjunto de demanda atribuído a facilidade x foi colorido com uma cor diferente, de forma a identificar as fronteiras de dados de cada facilidade.

No MATLAB, foi utilizado do algoritmo já implementado do *k-means*. Cada região de *cluster* retornado pelo método foi plotado com uma cor diferente para possibilitar a visualização das abrangências de cada *cluster*. Para o problema do posicionamento e cobertura das viaturas policiais, foi utilizado os pontos referentes aos centroides de cada *cluster*, retornados também pelo método supracitado. Estes dados podem ser observados e analisados na Figura 4.

Pode ser observado que a resposta obtida pelo método inteiro implementado no *CPLEX* valida o resultado obtido pelo *k-means*, já que ambos os métodos obtiveram resultados bem próximos. As saídas de ambos os métodos estão descritas em ordem na Tabela 1, onde foi obtido os endereços de centróide/facilidade.

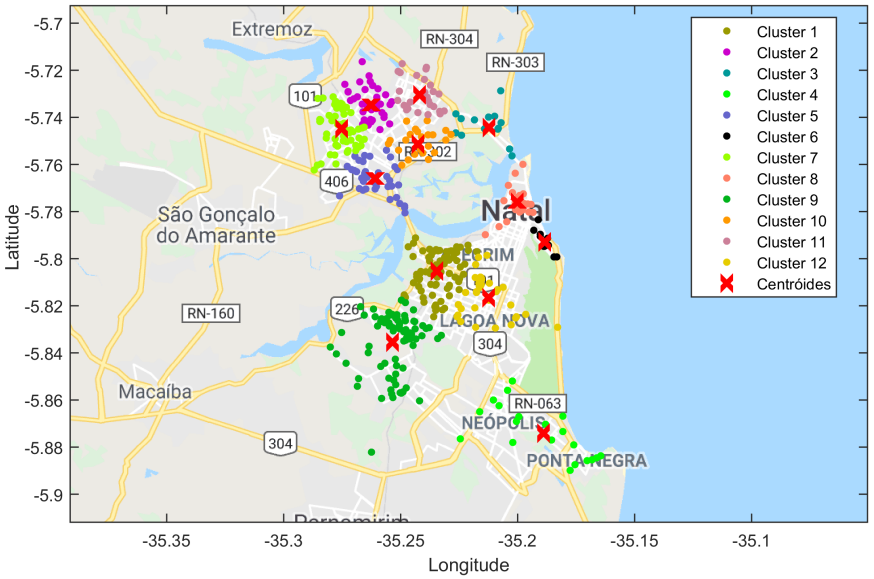
Figure 3: Mapa da distribuição de centroides gerados pelo modelo de *p-medianas* para as ocorrências de homicídios na cidade de Natal-RN.



Fonte: Autoria própria (2019)

[4]MATLAB: https://www.mathworks.com/products/matlab.html

Figure 4: Mapa da distribuição de centroides gerados pelo modelo de *k-means* para as ocorrências de homicídios na cidade de Natal-RN.



Fonte: Autoria própria (2019)

Table 1: Pontos de referências dos centroides gerados pelo modelo inteiro (P-medianas) e *K-means*

Modelo Inteiro	***K-means***
R. Santa Cristina, 382-a - Felipe Camarão, Natal - RN, 59072-150	R. Sagrada Família, 68 - Nossa Sra. da Apresentação, 59114-239
R. Henrique Dias, 1128 - Igapó, 59104-300	R. da Sanfona, 1081-99 - Lagoa Azul, 59138-420
Av. das Cirandas, 1001-1011 - Lagoa Azul, 59138-600	Av. Acaraú, 540 - Potengi, 59108-005
R. Jaguarari, 1848-1850 - Candelária, 59052-700	R. Curitiba, 377 - Potengi, 59124-440
Rua Iguatú, 370 - Potengi, 59124-010	R. dos Pioneiros, 698 - Pajuçara, 59125-340
Av. Dos Expedicionários - Conj. Pq. dos Coqueiros - Nossa Sra. da Apresentação, 59114-246	R. Padre Cícero Romão, 539 - Redinha, 59122-210
Av. Remador Clodoaldo Bakker, 884 - Pajuçara, 59125-300	R. Gen. Glicério, 851 - Rocas, 59012-100
R. Paracati, 3 - Planalto, 59073-100	R. Camaragibe, 665-635 - Mãe Luíza, 59014-330
Av. Praia de Muriú, 867-785 - Ponta Negra, 59092-390	Av. Nascimento de Castro, 1842 - Lagoa Nova, 59056-450
R. Altamira, 185-1 - Praia do Meio, 59010-120	4a Travess Antonio Basilio, 5 - Quintas, 59050-205
R. Augusto Leopoldo, 130-152 - Quintas, 59050-030	R. Prof. Antônio Trigueiro, 9 - Guarapes, 59074-607
Av. Ulisses Guimarães, 341-285 - Redinha, 59122-170	Av. Praia de Muriú, 284-310 - Ponta Negra, 59094-330

Fonte: Autoria própria (2019)

5. Considerações Finais

Com os resultados obtidos e demonstrados nesse estudo, é possível sugerir as autoridades competentes um reposicionamento de viaturas, seja por triangulação estratégica dos pontos georreferenciados obtidos pelos métodos, ou por outros meios. Uma possibilidade é utilizar dos dados por intervalos de tempo, seja mensalmente, trimestralmente ou semestralmente, possibilitando que as viaturas façam a cobertura das áreas violentas da cidade para inibir a ação criminosa.

Sugere-se como trabalhos futuros a validação por meio da comparação dos resultados obtidos uma meta-heurística. Como também utilizar de métodos de caminho mínimo para comparar as melhores rotas policiais entre os centroides/facilidades encontrados e os pontos de maiores incidências nos agrupamentos. Esse estudo pode ser complementado com uma análise temporal da criminalidade da cidade de Natal-RN. Fazendo a análise e a aplicação dos métodos utilizados nesse artigo para dados de períodos de tempo maiores, como 3, 5 e 7 anos. Com isso, pode ser possível realizar inferências um pouco mais precisas com o auxílio da análise estatística obtida nesse estudo.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

References

Arakaki, R. G. I. e Lorena, L. A. N. (2006). Uma heurística de localização-alocação (hla) para problemas de localização de facilidades. *Produção*, 16(2):319–328.

BBC (2018). Estas são as 50 cidades mais violentas do mundo (e 17 estão no brasil). URL `https://www.bbc.com/portuguese/brasil-43309946`. [Acessado: 2018-11-28].

Bezerra, S. N., Souza, S., e Almeida Vitor, J. (2008). Algoritmos evolutivos paralelos aplicados ao problema das p-medianas. *XL Simpósio Brasileiro de Pesquisa Operacional, João Pessoa, Paraíba, Brasil*.

BRASIL (1988). Constituição da república federativa brasileira. URL `http://www.planalto.gov.br/ccivil_03/constituicao/constituiÃğÃčo.htm`. [Acessado: 2019-06-04].

CCSPJP (2019). Metodologia del ranking (2018) de las 50 ciudades mas violentas del mundo. p. 43.

da Silva, V. F., Cabral, L. d. A. F., e Quirno, R. (2013). Estratégia para seleção de locais candidatos voltada ao problema de localização de viaturas policiais. *XLVSBPO-Simpósio Brasileiro de Pesquisa Operacional, Natal*.

Dantrakul, S., Likasiri, C., e Pongvuthithum, R. (2014). Applied p-median and p-center algorithms for facility location problems. *Expert Systems with Applications*, 41(8):3596–3604.

de Almeida, I. I. (2014). MetaheurÍstica hÍbrida utilizando grasp reativo e aprendizagem por reforÇo: Uma aplicaÇÃo na seguranÇa pÚblica.

de Freitas Neto, F. P., de Castro, A. F., e de Lima, F. d. C. (2014). Business intelligence applied in decision support for public safety. In *2014 9th Iberian Conference on Information Systems and Technologies (CISTI)*, p. 1–6. IEEE.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., e Tatham, R. L. (2009). *Análise multivariada de dados*. Bookman Editora.

IPEA (2017). Atlas da violência 2017 mapeia os homicídios no brasil. `http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=30253/`. Acessado: 2018-11-28.

Likas, A., Vlassis, N., e Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition*, 36(2):451–461.

Nasraoui, O. e N'Cir, C.-E. B. (2018). *Clustering Methods for Big Data Analytics: Techniques, Toolboxes and Applications*. Springer.

Pereira, M. A. (2005). Um método branch-and-price para problemas de localização de p-medianas. *São José dos Campos*.

Rabie, H. M., El-Khodary, I. A., e Tharwat, A. A. (2014). Particle swarm optimization algorithm for the continuous p-median location problems. In *2014 10th International Computer Engineering Conference (ICENCO)*, p. 81–86. IEEE.

Tan, P.-N., Steinbach, M., Kumar, V., et al. (2006). Cluster analysis: basic concepts and algorithms. *Introduction to data mining*, 8:487–568.

Tsihrintzis, G. A., Sotiropoulos, D. N., e Jain, L. C. (2019). Machine learning paradigms: Advances in data analytics. In *Machine Learning Paradigms*, p. 1–4. Springer.

Vicini, L. e Souza, A. M. (2005). Análise multivariada da teoria à prática. *Santa Maria: UFSM, CCNE*, p. 32.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings