



UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO ACADÊMICO DO AGRESTE

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

BRUNO FERREIRA DA COSTA BORBA

**PROPOSTA DE UM SISTEMA DE INFORMAÇÃO GERENCIAL PARA ANÁLISE
DE DADOS BASEADO NO MODELO K-MEANS E MCLP SOBRE A
LOCALIZAÇÃO DE INSTALAÇÕES POLICIAIS**

Caruaru

2019

BRUNO FERREIRA DA COSTA BORBA

**PROPOSTA DE UM SISTEMA DE INFORMAÇÃO GERENCIAL PARA ANÁLISE
DE DADOS BASEADO NO MODELO K-MEANS E MCLP SOBRE A
LOCALIZAÇÃO DE INSTALAÇÕES POLICIAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Engenharia de Produção.

Área de Concentração: Otimização e Gestão da Produção.

Orientadora: Prof.^a Dr.^a Ana Paula Henriques Gusmão de Araújo Lima.

Caruaru

2019

Catálogo na fonte:
Bibliotecário – Raul César de Melo - CRB/4 - 1735

B726p Borba, Bruno Ferreira da Costa.
 Proposta de um sistema de informação gerencial para análise de dados baseado no modelo K-MEANS e MCLP sobre a localização de instalações policiais. / Bruno Ferreira da Costa Borba. – 2019.
 117 f. ; il. : 30 cm.

 Orientadora: Ana Paula Henriques Gusmão de Araújo Lima.
 Dissertação (Mestrado) – Universidade Federal de Pernambuco, CAA, Programa de Pós-Graduação em Engenharia de Produção, 2019.
 Inclui Referências.

 1. Sistemas de informação gerencial. 2. Segurança pública. 3. Pernambuco. Polícia Militar - Instalações. I. Lima, Ana Paula Henriques Gusmão de Araújo (Orientadora). II. Título.

CDD 658.5 (23. ed.)

UFPE (CAA 2019-281)

BRUNO FERREIRA DA COSTA BORBA

**PROPOSTA DE UM SISTEMA DE INFORMAÇÃO GERENCIAL PARA ANÁLISE
DE DADOS BASEADO NO MODELO K-MEANS E MCLP SOBRE A
LOCALIZAÇÃO DE INSTALAÇÕES POLICIAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Engenharia de Produção.

Aprovada em: 04/12/2019.

BANCA EXAMINADORA

Prof.^a Dr.^a Ana Paula Henriques Gusmão de Araújo Lima (Orientadora)
Universidade Federal de Pernambuco

Prof.^a Dr.^a Thárcylla Rebecca Negreiros Clemente (Examinadora Interna)
Universidade Federal de Pernambuco

Prof. Dr. ° Thyago Celso Cavalcante Nepomuceno (Examinador Externo)
Universidade Federal de Pernambuco

Dedico a todos que sentiram minha ausência e tiveram a compreensão, que eu só não estava presente, porque estava estudando para alcançar essa vitória.

AGRADECIMENTOS

A meu Deus, pelas oportunidades e pela força concedida em meu coração nos momentos mais difíceis, permitindo meu crescimento para servi-lo da melhor forma.

A meus pais (João e Betânia), ao núcleo familiar da minha irmã (Valéria, Enio, Alice e Júlia) e a minha tia (Silvana) por ser meu refúgio e mesmo que distantes, estavam todos sempre rezando por mim e se esforçando ao máximo para possibilitar o realizar desse sonho.

A minha namorada Roberta Amorim, que ao longo desses anos sempre foi muito carinhosa e paciente em todos os momentos e principalmente nos dias que não foi possível o nosso encontro, mediante aos dias intensivos de estudo.

A minha orientadora Ana Paula, que fez parte do meu crescimento desde o início da minha graduação. E que agora no mestrado, sempre esteve próxima durante toda a execução deste trabalho, não importando a hora do dia da noite ou da madrugada, ela estava sempre presente ministrando os melhores caminhos a se seguir.

Ao grupo de estudos “Mira no A” (Saulo, Almiro, Dallas e Déborah), que sempre estava ligada em 220V a qualquer hora do dia e da noite para solucionar qualquer problema desse mestrado. Nunca vou esquecer de cada momento vivenciado e de toda ajuda a mim prestada. Principalmente por Saulo e Dallas, que me fizeram subir de nível com relação ao processamento de dados, sempre me incentivando e ajudando com meus problemas de programação no *RStudio*.

A família que foi gerada e fortalecida ao longo desse tempo por meus amigos Lucas e Vinicius, que dividiram apartamento comigo e foram meus amigos de todas as horas e que considero para o resto de minha vida. Onde esse sentimento se estende equivalentemente para Wilma e Vininho, que forneceram um grande suporte nos momentos crucias dessa jornada.

A todos meus professores que contribuíram para meu crescimento. Em especial ao Professor Gilson Lima e a Professora Marcele Fontana, que foram meus orientadores em outro ponto do tempo e deixaram sua marca no meu jeito de trabalhar.

Ao meu amigo George Andrade, que estava na secretaria do mestrado e sempre esteve disponível para ajudar. E a todos que contribuíram, direta ou indiretamente, para elaboração deste trabalho e para conclusão do meu curso, os meus sinceros agradecimentos.

“Essa vida é uma ilusão quem quiser que se iluda, estamos aqui para ser feliz!” (COSTA e VASCONCELOS, 2000, p.28)

RESUMO

Está cada vez mais notório para a sociedade que os índices de criminalidades no Brasil vêm aumentando muito nos últimos anos. Tal fato provoca uma sensação de medo na população e principalmente naqueles que vivem em grandes centros urbanos. A região Nordeste do Brasil apresenta-se em segundo lugar no ranking de criminalidade, na qual, o estado de Pernambuco, nos últimos dois anos, teve aumentos consideráveis em seus investimentos voltados a segurança pública e seus índices de criminalidade não reduziram. O que chama a atenção dos gestores responsáveis é que, por se tratar de recursos públicos, é imperativo que todos os esforços sejam voltados a maximizar o retorno desses investimentos. Porém, inúmeros fatores dificultam a decisão de alocação desses recursos, sendo um dos principais fatores a falta de sistemas de suporte gerencial a decisão. Nesse sentido, sabendo que a localização das instalações policiais influencia de forma positiva nos índices de criminalidade, especificamente no tempo de resposta policial, esta pesquisa tem como objetivo propor um Sistema de Informação Gerencial (SIG) que apresente diferentes arranjos de localizações, com seus respectivos indicadores de desempenho e que, dessa forma, suporte decisões a respeito da instalação de unidades policiais na cidade do Recife-PE, de forma a aumentar a eficiência operacional no atendimento às ocorrências policiais e no patrulhamento nas regiões cobertas. Logo, a metodologia utilizada no SIG, para localizar de forma otimizada as instalações policiais, faz uso de dois modelos, o primeiro voltado para encontrar os locais candidatos, chamado de *K-means*, e o segundo voltado a selecionar os locais ótimos, chamado de *Maximal Covering Location Problem (MCLP)*. Posteriormente ao desenvolvimento do SIG, o mesmo foi aplicado com base nos dados da plataforma “Onde Fui Roubado” referente às ocorrências da cidade de Recife-PE, obtendo assim, resultados melhores quando comparado com a atual localização das unidades dessa cidade. Sendo assim, ao fim desta pesquisa foi concluído que o SIG proposto atingiu com sucesso o objetivo inicialmente da pesquisa, desenvolvendo uma ferramenta inovadora para a modelagem e resolução do problema de localização de instalações policiais.

Palavras-chave: MCLP. *K-mean*. Sistema de Informação Gerencial. Localização de Instalações. Segurança Pública.

ABSTRACT

It is becoming increasingly apparent to society that crime rates in Brazil have been increasing sharply in recent years. This causes a sense of fear in the population and especially those living in large urban centers. The Northeast region of Brazil ranks second in the crime ranking, in which the state of Pernambuco, over the last two years, has had considerable increases in its investments in public security and its crime rates have not reduced. What draws the attention of the managers responsible is that, as these are public resources, it is imperative that all efforts are directed at maximizing the return on these investments. However, many factors make it difficult to decide on the allocation of these resources, one of the main factors being the lack of decision support management systems. Knowing that the location of police facilities positively influences crime rates, specifically police response time, this research aims to propose a Management Information System (GIS) that presents different location arrangements, with their respective performance indicators and, thus, support decisions regarding the installation of police units in the city of Recife-PE, in order to increase operational efficiency in responding to police occurrences and patrolling in the covered regions. Therefore, the methodology used in the GIS to optimally locate police facilities makes use of two models, the first one aimed at finding candidate sites, called K-means, and the second one looking at selecting optimal locations, called Maximal Covering Location Problem (MCLP). Subsequent to the development of the GIS, it was applied based on data from the platform “Onde Fui Roubado” referring to occurrences of the city of Recife-PE, thus obtaining better results when compared to the current location of the units of this city. Thus, at the end of this research it was concluded that the proposed GIS successfully achieved the initial objective of the research by developing an innovative tool for modeling and solving the problem of police facility location.

Keywords: MCLP. K-Means. Management Information System. Facilities Location. Public security.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Taxa de mortes por 100 mil habitantes em 2017	23
Gráfico 1 –	Número e da taxa de homicídios entre 2006 a 2016	23
Gráfico 2 –	Taxa de homicídios entre 2006 a 2016 por região	24
Figura 2 –	Ilustração de localização pelo P-median problem.....	31
Figura 3 –	Ilustração de localização pelo Center Problems.....	33
Figura 4 –	Ilustração de localização pelo Set Covering Problem.....	36
Figura 5 –	Ilustração de localização pelo Maximal Covering Location Problem	38
Figura 6 –	Conjunto de demanda categorizado empiricamente em clusters	42
Figura 7 –	Ilustração dos passos do método k-means	44
Figura 8 –	Arquitetura básica do Sistema de Informação Gerencial.....	47
Fluxograma 1 –	Processo de revisão da literatura	48
Quadro 1 –	Palavras chaves e peradores.....	50
Gráfico 3 –	Publicação anual científica sobre o modelo MCLP	51
Gráfico 4 –	Fontes analisadas na revisão literária	53
Gráfico 5 –	Relação de publicações por autor	53
Gráfico 6 –	Produção dos autores ao longo do tempo.....	54
Gráfico 7 –	Referências mais citadas	55
Figura 9 –	Produção de documentos por país.....	56
Fluxograma 2 –	Delineamento da pesquisa.....	59
Fluxograma 3 –	Implementação do modelo MCLP no Software RStudio.....	61
Figura 10 –	Arquitetura do SIG proposto.....	64
Figura 11 –	Página Arquivo do SIG proposto	67
Figura 12 –	Página Dados do SIG proposto, guia Dados Estatísticos.....	68
Figura 13 –	Página Dados do SIG proposto, guia Mapa Crimes.....	69
Figura 14 –	Página Kmeans do SIG proposto, guia Kmeans	70
Figura 15 –	Página Kmeans do SIG proposto, guia Mapa Cluster 1	71
Figura 16 –	Página Kmeans do SIG proposto, guia Mapa Cluster 2.....	71
Figura 17 –	Página Locais Candidatos do SIG proposto, desvio	72
Figura 18 –	Página Locais Candidatos do SIG proposto, cobertura.....	73
Figura 19 –	Página Locais Candidatos do SIG proposto, não coberto	74
Figura 20 –	Página Locais Candidatos do SIG proposto, relatório	75
Figura 21 –	Página Alocação (MCLP) do SIG proposto, quantidade	76

Figura 22 –	Página Alocação (MCLP) do SIG proposto, locais selecionados	76
Figura 23 –	Página Alocação (MCLP) do SIG proposto, raio.....	77
Figura 24 –	Página Download do SIG proposto.....	78
Gráfico 8 –	Relação de Crimes por Ano.....	82
Gráfico 9 –	Quantitativo de Crimes por Tipo.....	82
Gráfico 10 –	Percentual que Registrou Boletim de Ocorrência.....	83
Gráfico 11 –	Percentual de Vítimas por Gênero.....	83
Mapa 1 –	Mapa de Crimes em Recife	84
Gráfico 12 –	Bairros Com Mais Ocorrência de Crimes em Recife	85
Mapa 2 –	Clusters de Crimes na Cidade do Recife	86
Quadro 2 –	Coordenadas dos locais candidatos	86
Mapa 3 –	Cluster Distância	87
Mapa 4 –	Georreferência dos Locais Candidatos	87
Quadro 3 –	Cobertura ótima para cinco instalações alocadas	88
Mapa 5 –	Instalações ótimas alocadas, com raio de 3.0 km de cobertura	89
Gráfico 13 –	Incremento de cobertura por instalações, com raio de 3km.....	90
Gráfico 14 –	Redução da distância média por instalações, com raio de 3km	91
Gráfico 15 –	Comparação de desempenho, com relação a cobertura.....	92
Gráfico 16 –	Comparação de desempenho, com relação a distância média.....	93
Mapa 6 –	Instalações ótimas alocadas, com raio de 1.0 km de cobertura	112
Mapa 7 –	Instalações ótimas alocadas, com raio de 1.5 km de cobertura	112
Mapa 8 –	Instalações ótimas alocadas, com raio de 2.0 km de cobertura	113
Mapa 9 –	Instalações ótimas alocadas, com raio de 2.5 km de cobertura	113
Mapa 10 –	Instalações ótimas alocadas, com raio de 3.5 km de cobertura	114
Mapa 11 –	Instalações ótimas alocadas, com raio de 4.0 km de cobertura	114
Mapa 12 –	Instalações ótimas alocadas, com raio de 4.5 km de cobertura	115
Mapa 13 –	Instalações ótimas alocadas, com raio de 5.0 km de cobertura	115
Gráfico 17 –	Análise com a quantidade de instalações fixas e raio variável.....	116
Gráfico 18 –	Análise com o raio fixo e instalações variáveis	117

LISTA DE SIGLAS

MCLP	<i>Maximal Covering Location Problem.</i>
CVLI	Crimes Violentos Letais Intencionais.
PO	Pesquisa Operacional
DM	Decisor
MS	Ministério da Saúde
SDS	Secretaria de Defesa Social
SCP	<i>Set Covering Problem</i>
HFLP	<i>Hierarchical facility location problem</i>
HLP	<i>Hub location problems</i>
IBM	<i>International Business Machine</i>
GA	<i>Greedy Adding</i>
GAS	<i>Greedy Adding with Substitution</i>
MPS	Sistema de Programação Matemático
GIS	Sistema de Informação Geográfica
ISI	<i>Institute for Scientific Information</i>
MCLPDC	<i>Minimum Weighted Covering Location Problem With Distance Constraints.</i>
GPSID	Grupo de Pesquisa em Sistemas de Informação e Decisão.
UFPE	Universidade Federal de Pernambuco.
SIG	Sistema de Informação Gerencial.
BO	Boletim de Ocorrência.
IBGE	Instituto Brasileiro de Geográfica e Estatística
PM	Polícia Militar

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Justificativa	17
1.2	Objetivo Geral.....	18
1.3	Objetivos Específicos.....	18
1.4	Estruturação do trabalho.....	19
2	REFERENCIAL TEÓRICO	21
2.1	Segurança Pública	21
2.1.1	Cenário da segurança pública no Brasil	22
2.1.2	Região Nordeste e o estado de Pernambuco.....	24
2.2	Problema de localização de Instalações	25
2.2.1	Diferenças na localização de instalações públicas e privadas	27
2.2.2	Localização de instalações voltadas a segurança pública.....	28
2.3	Modelos voltados a localização de instalações	29
2.3.1	P-median problem.....	30
2.3.2	Center problems.....	32
2.3.3	Covering problems	34
2.3.3.1	<i>Set Covering Problem.....</i>	<i>35</i>
2.3.3.2	<i>Maximum Covering Location problem.....</i>	<i>36</i>
2.3.4	Outros modelos abordados na literatura	40
2.4	Clusters como inputs para modelos de localização.....	41
2.4.1	Método K-means	42
2.5	Sistema de informação gerencial.....	45
3	REVISÃO DA LITERATURA NO MODELO MCLP	48
3.1	Método da Revisão.....	48
3.1.1	Planejamento da revisão	49
3.1.2	Conduta da revisão	49

3.1.3	Documento de revisão	51
3.2	Justificativa do uso do modelo.....	56
4	<i>METODOLOGIA</i>	58
4.1	Classificação da pesquisa científica em engenharia de produção	58
4.2	Delineamento com as fases da pesquisa	58
4.2.1	Definição do local da pesquisa	59
4.2.2	Identificação da problemática.....	59
4.2.3	Estudo da literatura sobre a problemática.....	60
4.2.4	Coleta e tratamento dos dados	60
4.2.5	Desenvolvimento do SIG.....	61
4.2.6	Análise de sensibilidade	62
4.2.7	Análise dos resultados	63
5	<i>SISTEMA DE INFORMAÇÃO GERENCIAL PROPOSTO</i>	64
5.1	Arquitetura do SIG	64
5.2	Design e funcionamento do SIG proposto	66
6	<i>ESTUDO DE CASO</i>	80
6.1	Local de estudo.....	80
6.2	Base de dados	80
6.3	Dados Iniciais	81
6.4	Locais Candidatos.....	85
6.5	Alocação ótima.....	88
6.6	Comparação do cenário atualmente utilizado com o proposto	92
7	<i>CONSIDERAÇÕES FINAIS</i>	95
7.1	Contribuições da pesquisa	95
7.2	Dificuldades e limitações	96
7.3	Trabalhos futuros	96
	<i>REFERÊNCIAS</i>	99

<i>APÊNDICE A – RESULTADOS DA ALOCAÇÃO OTIMIZADA</i>	<i>112</i>
<i>APÊNDICE B – ANÁLISE DE SENSIBILIDADE</i>	<i>116</i>

1 INTRODUÇÃO

Nos últimos anos, ficou notório o crescimento dos índices de criminalidades no Brasil e nos demais países da América Latina (TELLA *et al.*, 2010; CERQUEIRA *et al.*, 2007). Esse rápido crescimento dos índices de criminalidade geram uma sensação de medo na população, principalmente dos que convivem em grandes capitais, no qual, o nível de violência tem se apresentado em maiores números (SOHNGEN e CIPRIANI, 2019).

No Brasil, esses índices de criminalidades são ainda mais “alarmantes” na região Nordeste. Segundo Nóbrega Jr. (2015) a região Nordeste foi responsável por 37% dos homicídios totais do país, apresentando um incremento percentual de 89% no comparativo 2000/2011. Onde foi observado, um aumento significativo no estado de Pernambuco nos últimos 4 anos para os CVLI (Crimes Violentos Letais Intencionais). Esse aumento de CVLI em Pernambuco, consta de 3434, 3889, 4480 e 5426 vítimas, para os respectivos anos de 2014, 2015, 2016 e 2017 (10ª ANUÁRIO DE SEGURANÇA PÚBLICA, 2016; 11º ANUÁRIO DE SEGURANÇA PÚBLICA, 2017; 12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018).

Apesar dos aumentos com relação a criminalidade, houve um aumento com relação aos investimentos voltados a segurança pública de Pernambuco. Esses saíram de R\$ 2.394.482.512,88 em 2016 para R\$ 2.529.880.338,52 em 2017, ponto de reflexão para os tomadores de decisão responsáveis, que mesmo investindo mais, não se observa redução nos índices de criminalidade. Principalmente pelo maior aumento entre as três categorias (Policimento, Defesa Civil e Informação e Inteligência) ter sido efetuado em Informação e Inteligência, representando uma variação de 97,49% com relação ao ano anterior (12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018).

Esse aumento dos índices de criminalidade e dos investimentos é alarmante e logo acaba se tornando objeto de estudo para diversas áreas do conhecimento. Onde o principal foco é melhor entender o que acontece para que os mesmos ocorram e para se reduzir ao máximo esses indicadores. O fato é que o crime e o medo do crime, acabam se tornando um grave risco para a população e devem ser gerenciados e estimados para mitigá-los e assim fornecer um ambiente melhor para a população (GARLAND, 2014).

Visando a minimização desses índices criminais e a maximização dos recursos policiais, surgiu o interesse dos principais executivos e tomadores de decisão na utilização de técnicas de Pesquisa Operacional (PO) (MITCHELL, 1972; GUSMÃO *et al.*, 2019). E uma das áreas de

estudos que vem se mostrando bastante promissora ao longo do tempo com relação à segurança pública, é localização ótima de unidades policiais voltadas a segurança e preservação da ordem (GURGEL, 2010; PIZZOLATO, 2012; LORENA, 2003; REVELLE, 2005; GUSMÃO *et al.*, 2019).

Nesse sentido, segundo Owen e Daskin (1998) a construção ou compra de novas instalações policiais é tipicamente um projeto caro e sensível ao longo do tempo. Antes da execução do projeto é preciso analisar se as localizações das unidades policiais são viáveis mediante a seus objetivos operacionais e estratégicos. A PO surge para ajudar os tomadores de decisão nessa difícil tarefa de escolher um local que não somente irá resultar em um melhor desempenho considerando a estrutura atual, mas que seja promissora no futuro, mesmo que os fatores ambientais mudem ou evoluam (FALLAH, 2009).

Logo, diversos estudos nessa área foram realizados e conseguiram melhorar o desempenho de forma significativa, no qual, o cenário se encontrava. Esses estudos iniciaram na década de 60, quando se buscava encontrar um modo de minimizar a distância média entre as instalações e a demanda (HAKIMI, 1964). Posteriormente outros autores trabalharam em localizar instalações emergenciais (TOREGAS *et al.*, 1971), como localizar duas novas estações policiais para suplementar uma sede da polícia já existente (ALY; LITWHILER, 1979).

Outro trabalho que representou uma grande contribuição no contexto da PO aplicada à segurança, e que se adequa ao objetivo de maximizar a atuação policial mediante aos seus recursos, é escrito por Church e Reville na década de 70. Nesse trabalho, foi formulado um modelo chamado de *Maximal Covering Location Problem* (MCLP) em que se procura maximizar a cobertura dentro de uma distância de serviço desejada, localizando um número fixo de instalações (CHURCH e REVELLE, 1974).

Entretanto algumas diferenças podem ocorrer quando é preciso alocar uma instalação pública. Cirino *et al.* (2018) afirmam que diferente das instalações privadas que buscam maximizar seus lucros, as instalações públicas, e principalmente as de caráter emergencial visam estar o mais próximo possível da população para garantir seu bem-estar e segurança, algo que não é quantificável em dinheiro. Dessa forma, será possível encaminhar qualquer tipo de serviço emergencial para o ponto de demanda em questão, em tempo hábil, principalmente quando se trata de serviços policiais.

Em virtude desses fatos são notórias as dificuldades que são enfrentadas (tempo de resposta para a ocorrência, local da instalação, área de cobertura, quantidade de instalações) pelo tomador de decisão, dado que o mesmo deve levar em consideração diversos fatores para maximizar o retorno de seu investimento no momento de alocar uma instalação policial. Considerando o exposto, este trabalho propõe o desenvolvimento de um sistema de informação gerencial (SIG) para suportar questões gerenciais e operacionais relativas à localização de instalações policiais, visando identificar alternativas de localização que maximizem o retorno dos investimentos em função da qualidade do serviço prestado.

1.1 Justificativa

Mediante o cenário vivenciado nos últimos anos no estado de Pernambuco, no qual foi observado um aumento gradual dos investimentos em departamentos públicos que são voltados para manter a ordem e a segurança da população. Fez com que os tomadores de decisões refletissem sobre os seus investimentos, dado que, os índices criminais continuam aumentando e que em algum momento não será mais possível incrementar no valor do capital que já é investido.

Entretanto, diversos fatores influenciam essa questão e nem sempre é possível obter os melhores resultados. E por se tratar de recursos públicos, é imperativo que todos os esforços sejam voltados em maximizar a alocação desses investimentos em pontos que são mais urgentes e que gerem mais benefícios para a população.

Nesse sentido, em muitos dos casos a grande dificuldade em realizar a alocação correta dos investimentos está relacionada à falta de um sistema de informação gerencial que venha dar um suporte formal à tomada de decisão. Suporte esse que faz uso de um conjunto de bancos de dados e modelos para o processamento e desenvolvimento de diretrizes que facilitem na identificação da melhor alternativa de ação.

Logo, como o investimento da localização de uma instalação policial tem impactos diretos nos índices de criminalidades de uma região de forma positiva, faz sentido a utilização de modelos formais de alocação de instalações. Ou seja, busca-se alocar instalações policiais em pontos que maximizem a atuação policial mediante os recursos que o departamento de segurança pública tem disponível.

Sabendo qual é a melhor localização para se investir em uma instalação policial, faz com que os esforços sejam direcionados, aumentando a eficiência da segurança pública e até podendo reduzir os gastos desnecessários. Quanto mais próximas estiverem as instalações da

população que demanda os serviços policiais, maior será a quantidade de pessoas cobertas e menos tempo será gasto entre a instalação receber o chamado e enviar uma viatura ao local da ocorrência.

Considerando o volume de dados a serem processados e a velocidade de resposta requeridas pelos decisores, torna-se notória a necessidade de mais estudos para o desenvolvimento de um sistema de informação gerencial, visto que é necessário mitigar essas divergências entre os investimentos em segurança pública e os índices de criminalidade. Tendo como um ponto inicial, o problema da localização das instalações policiais, que posteriormente pode ser realizado análises entre as localizações das instalações e as regiões que estão acontecendo mais crimes, com objetivo de encontrar pontos que estejam fora de conformidade perante o escopo da polícia.

1.2 Objetivo Geral

Esta pesquisa tem como objetivo geral, propor um sistema de informação gerencial (SIG), baseado em modelos e técnicas computacionais, tais como o K-means e o MCLP, para dar suporte a análise e visualização de dados sobre o arranjo de instalações policiais na cidade do Recife-PE.

No próximo tópico serão abordados os objetivos específicos necessários para o alcance deste objetivo geral definido.

1.3 Objetivos Específicos

Para atender a este objetivo geral supracitado, os seguintes objetivos específicos devem ser atingidos:

- Realizar uma pesquisa exploratória sobre o modelo MCLP e sobre os temas relacionados a sua problemática;
- Delimitar uma área/região que será objeto de estudo e para a qual se disponha de índices de criminalidades;
- Desenvolver e explorar a aplicabilidade do SIG proposto na área de estudo e compará-lo com o método existente;
- Realizar análises e discussões dos resultados obtidos.

1.4 Estruturação do trabalho

A estrutura da dissertação é apresentada com o objetivo de facilitar o entendimento dos conceitos que são revelados mediante o avançar dos estudos. Nesse sentido, a dissertação apresenta-se estruturada em sete capítulos, cujos assuntos explanados em cada capítulo serão abordados a seguir:

- Capítulo 1: são apresentados os aspectos iniciais da pesquisa, contendo uma introdução que sintetiza o tema a ser explorado, ou seja, os aumentos dos índices criminais no Brasil e a problemática de localização das instalações. Seguindo com a justificativa, e por fim os objetivos geral e específicos deste trabalho;
- Capítulo 2: é apresentado o referencial teórico, no qual se aborda os principais conceitos para o entendimento do estudo. Iniciando com os conceitos sobre segurança pública, problema de localização de instalações, modelos voltados a localização de instalações, clusters como inputs para modelos de localização e por fim os conceitos voltados aos SIGs.
- Capítulo 3: é apresentada uma revisão da literatura sobre o modelo MCLP, assim como a metodologia utilizada para sua construção, no qual definem-se os limites da pesquisa, analisando sobre uma perspectiva científica e explorando todos os requisitos necessários para justificar o uso do modelo MCLP atualmente.
- O Capítulo 4: é apresentado a metodologia geral utilizada no estudo desta dissertação, ou seja, a diretriz seguida desde os estudos iniciais até a apresentação dos resultados. Define-se também, a caracterização e o delineamento da pesquisa de forma que o leitor possa entender melhor como foi realizada a construção desta pesquisa.
- O Capítulo 5: é apresentado a arquitetura do SIG proposto concomitantemente com design do SIG para solucionar o estudo de caso dessa pesquisa. Nesse sentido, toda a interação que o decisor pode ter com o sistema é apresentado através das telas do próprio SIG desenvolvido.
- O Capítulo 6: é apresentado o estudo de caso desenvolvido nesta pesquisa com base na aplicação do SIG proposto. Analisam-se e apresentam-se os resultados obtidos e compara-se com os resultados vivenciados na região nos dias atuais.
- O Capítulo 7: são apresentadas as considerações finais da pesquisa, incluindo os pontos fortes e fracos. Além da perspectiva para o desenvolvimento de trabalhos futuros.

Por fim, são apresentadas as referências utilizadas para a elaboração desta dissertação e os apêndices com alguns dos resultados obtidos durante a pesquisa.

2 REFERENCIAL TEÓRICO

Neste capítulo serão apresentados os principais conceitos, que foram essenciais para o desenvolvimento dessa dissertação. Inicialmente será apresentada uma breve descrição dos conceitos voltados a segurança pública. Logo em seguida serão trabalhados os conceitos sobre o problema de localização e as diferenças entre essa temática no setor público e privado. Seguindo para alguns dos principais modelos de localização de instalações, assim como, os clusters como inputs para o problema de localização de instalações e por fim os conceitos básicos dos sistemas de informações gerenciais.

2.1 Segurança Pública

O conjunto de conceitos voltados a políticas de segurança pública, nascido na modernidade, e que domina nos governos atuais, evidencia a estruturação do processo civilizacional. E nesse sentido, Carvalho e Silva (2011) procura demonstrar que a sociedade está constante movimento de altos e baixos, motivadas pela própria dinâmica que está em curso neste século, e que impõe a necessidade de segurança pública como garantia do exercício da cidadania.

Segundo Lima *et al.* (2016), o termo segurança pública é mencionado pela primeira vez na Constituição Federal de 1937, no qual, esse termo foi apresentado com outras palavras em versões anteriores da Constituições, como a de 1934, onde foi usado o termo segurança “interna”, para trabalhar com matérias referentes ao controle da ordem. A Constituição federal de 1988 afirma que a segurança pública é dever do estado, direito e responsabilidade de todos e é voltada para a preservação da ordem pública de forma que não exista perigo tanto para as pessoas como para os patrimônios, através dos seguintes órgãos de força-tarefa: polícia federal; polícia rodoviária federal; polícia ferroviária federal; polícias civis; polícias militares e corpos de bombeiros militares.

Nesse sentido, a segurança pública tem como objetivo principal, manter a ordem e o equilíbrio da sociedade visando o bem em conjunto, através dos poderes do estado e da cooperação da população. Assim a segurança pública previne o crime, enfrenta a problemática de medo da população, reduz os custos associados a questões de crimes, estabelece condições concretas e não meramente simbólicas, de segurança e são atividades que necessitam de responsabilidade de um conjunto de instituições fornecidas pelo governo, para garantir a segurança do território em questão (DIAS NETO, 2005).

Já Santos *et al.* (2015) complementam que, a segurança pública não deve ser compreendida apenas como uma atividade ou resultado do trabalho de instituições públicas, que fazem uso de seu poder legal para manter um controle social. Muito menos que, a segurança pública não deve ser levada em consideração a partir da individualização de causas ou de responsabilidades, porém pelo entendimento global das relações entre as causas e os efeitos atrelados aos processos sociais que o circundam (SILVA, 2017).

Nesse sentido, quando se trabalha com segurança pública é preciso analisar o cenário por etapas. Iniciando a partir de visão global do problema e restringindo para um ponto específico do problema. Assim, é possível estrutura-lo e levantar as relações envolvidas no problema para encontrar um conjunto de soluções que melhor se adequem.

2.1.1 Cenário da segurança pública no Brasil

O cenário no qual o Brasil vem enfrentando nos últimos 10 anos vem despertando atenção, tanto dos políticos como da sociedade para a temática de segurança pública, dado que a mesma vem se configurando como um assunto mais amplo e não somente, exclusivo das pautas corporativas de organizações policiais estaduais (SANTOS *et al.*, 2015). Esse foco na segurança pública se destaca devido ao aumento dos índices de criminalidade. Que são decorrentes de algumas falhas no sistema de segurança pública brasileira, que carece de amplas reformas administrativas e institucionais a fim de garantir espaços seguros de convivência democrática e promoção da cultura de paz (SOARES, 2006; SENTO-SÉ, 2006; CANO, 2006).

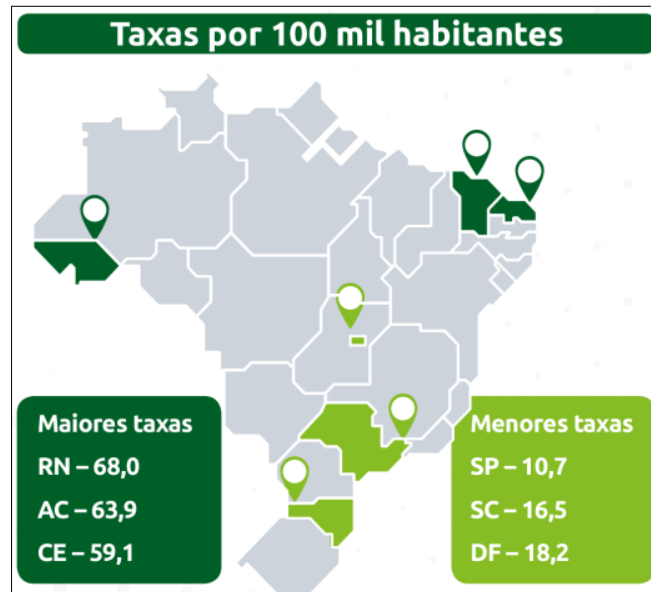
Corroborando Lima *et al.* (2015) afirmam que, a história recente da segurança pública no Brasil tem sido marcada por demandas acumuladas e mudanças incompletas em seu sistema. No qual, os eventuais avanços na gestão policial e reformas na legislação penal têm se mostrado insuficientes para reduzir os índices de criminalidade urbana, destacando a falta de ferramentas para a gestão e controle por parte dos órgãos responsáveis (LIMA *et al.*, 2016).

Essa situação pode ser comprovada por dados quantitativos levantados pelo Ministério da Saúde (MS), que destaca uma marca histórica no Brasil de 62.517 homicídios em 2016. Dados esse, que só pioram em 2017, chegando a 63.895 mortes violentas intencionais, são 175 homicídios por dia, atingindo uma taxa média 30,8 mortes por 100 mil habitantes (12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018).

Mediante a representação geográfica da Figura 1 retirada do 12º Anuário de Segurança Pública, é possível verificar quais estados Brasileiros tiveram as maiores e menores taxas de

mortes por 100 mil habitantes. Evidenciando as maiores taxas nos estados do Rio Grande do Norte, Acre e Ceará. E os com menores taxas em São Paulo, Santa Catarina e Distrito Federal.

Figura 1 – Taxa de mortes por 100 mil habitantes em 2017

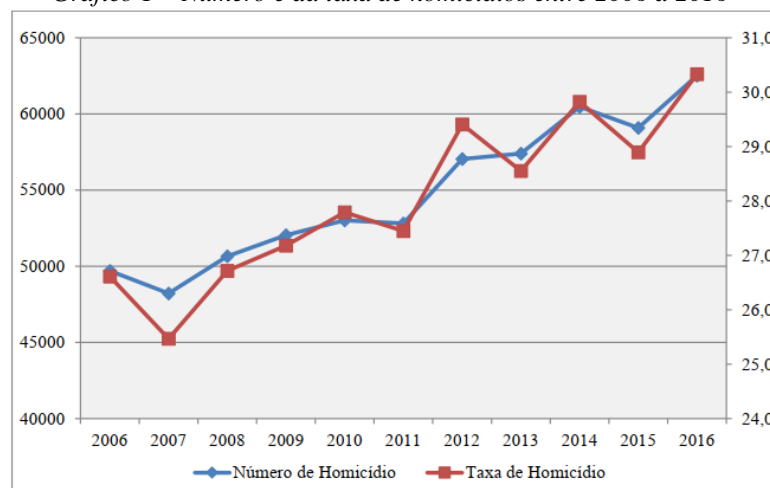


Fonte: 12º ANUÁRIO DE SEGURANÇA PÚBLICA (2018)

Os números já apresentam um grande impacto quando analisados desta forma. Porém, quando analisados em uma faixa histórica maior, acaba se tornando evidente o cenário da segurança pública no Brasil. Na Gráfico 1, retirada do Atlas da Violência de 2018 é apresentado o crescimento descontrolado do número e da taxa de homicídios entre 2006 a 2016 no Brasil.

Um claro crescimento que concretiza uma mudança no nível do indicador de homicídios, passado para o intervalo de 60 mil a 65 mil casos por ano. Segundo o IPEA (2018) esse indicador crescente relata, além da naturalização do fenômeno, a urgência que deve ser dada a essa problemática por parte das autoridades nos três níveis de governo: federal, estadual e municipal.

Gráfico 1 – Número e da taxa de homicídios entre 2006 a 2016



Fonte: IPEA (2018)

Segundo Trindade (2015) esse problema acontece por limitações e conflitos de competência entre os responsáveis nos três níveis de governo, que dificultam a implantação de novas políticas, além da cultura e estrutura interna das instituições de segurança pública e justiça criminal que dificultam no processo de inovação. Nesse sentido, essas e outras consequências relacionadas ao problema, podem trazer implicações na dinâmica demográfica e, por conseguinte, no processo de desenvolvimento econômico e social (IPEA, 2018).

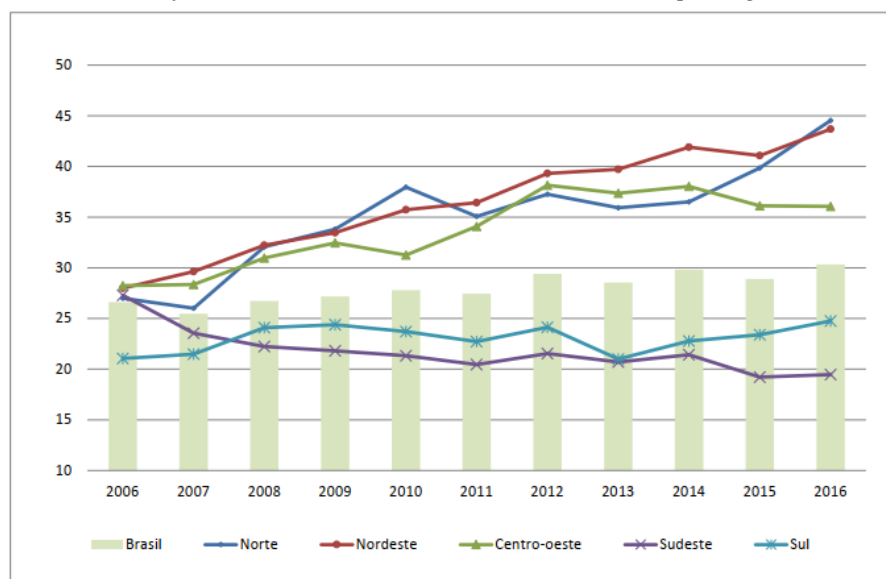
Quando são filtrados os dados em cada região do país, é possível perceber que é geral o aumento da criminalidade. Porém, a região Norte e Nordeste se destacam das demais.

2.1.2 Região Nordeste e o estado de Pernambuco

A região Nordeste vem apresentando altos níveis de crimes violentos, que se agravam mais em determinados estados e menos em outros. O Gráfico 2 retirado do Atlas da Violência de 2018, apresenta as taxas de homicídios por regiões entre intervalo de 2006 a 2016.

É notório o crescimento exorbitante da região Nordeste com relação a taxa de homicídios. Segundo o IPEA (2018) todos os estados com crescimento superior a 80% nas taxas de homicídios pertenciam ao Norte e ao Nordeste. Nóbrega Jr. e Zaverucha (2017) afirmam que essa problemática de crescimento constante de mortes por agressão nas Região Norte e Nordeste, vem de muito antes, desde 1996. O que confirma a necessidade urgente de reforma, nas políticas de segurança pública.

Gráfico 2 – Taxa de homicídios entre 2006 a 2016 por região



Fonte: IPEA (2018)

Com base no Banco de Dados do Subsistema de Informação de Mortalidade do Sistema Único de Saúde (SIM/Datasus), em Pernambuco durante o período de 2000 até 2013 houve

uma grande redução em seus índices de criminalidade. Essa melhoria no estado acontece a partir do ano de 2007 quando sucesso foi atingido no programa de contenção da violência (ZAUERUCHA e NÓBREGA JR., 2015).

Essa redução dos índices criminais em Pernambuco não durou muito e os índices voltaram a crescer descontroladamente. Segundo a Secretaria de Defesa Social de Pernambuco (SDS), de 2013 a 2016 houve um acréscimo de 44% da taxa de homicídios, e o que antes era 33,6 mortes por 100 mil habitantes, se tornaram 47 mortes por 100 mil habitantes. Essa situação se torna mais crítica em 2017, onde esses valores cresceram ainda mais, chegando a um total de 5426 vítimas em Pernambuco (12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018).

Segundo Nóbrega JR (2017), cidades que antes eram pacatas e tranquilas, agora se deparam com uma grave situação de criminalidade, citando como exemplo Petrolina e Caruaru. E a população que se encontra na capital, sofre com uma sensação de medo, mediante ao nível de violência que se apresentado em números maiores (SOHNGEN e CIPRIANI, 2019).

Logo, não pode ser ignorado os aumentos de investimentos que foram realizados nos últimos anos em Pernambuco, voltados a solucionar esse problema. Foram gastos em segurança pública 2,39 bilhões de reais em 2016 e 2,53 bilhões de reais em 2017 (12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018). Levanta preocupações, dado que no processo de criação de políticas em segurança pública, a busca pela eficiência e eficácia na utilização dos recursos públicos escassos, é um fator importante e que deve ser priorizado (GALDINO; GUIMARÃES; FILHO, 2018).

Segundo Waiselfisz (2014), é de extrema importância reexaminar urgentemente as práticas empregadas na segurança pública, buscando desenvolver e implementar novos modelos que melhor otimizem a aplicação dos investimentos em segurança. Portanto, métodos e técnicas de pesquisa operacional são primordiais para solução desse tipo de problema (GURGEL, 2010).

Visando essa importância da segurança pública para a população, se faz necessário cada vez mais, estudar como a mesma é desenvolvida e implementada. Apenas assim, será possível desenvolver novos métodos e metodologias para auxiliar e propor uma melhoria contínua de seus procedimentos, para garantir a segurança e a ordem.

2.2 Problema de localização de Instalações

Os estudos voltados a problemas de localização de instalações começaram de forma teórica em 1909, quando Alfred Weber teve a ideia de posicionar um único armazém de modo a minimizar a distância total entre os vários pontos de demanda (WEBER, 1929). Após esse

estudo, a teoria voltou a ter destaque em 1964, quando Hakimi buscou localizar centrais de comunicação em uma rede e delegacias de polícia dentro de um sistema rodoviário (HAKIMI, 1964).

Com base nesses percursos, o problema de localização de instalações (*Facility Location Problem*), consiste em alocar uma ou mais instalações em determinada localização atendendo um conjunto de critérios pré-estabelecidos. Em outras palavras, procura selecionar a melhor alternativa possível de localizar um dado número de instalações, de forma a servir um elevado nível de serviço aos clientes, minimizando os custos da operação e maximizando os seus lucros (MAPA e LIMA, 2012; FARAHANI et al., 2019).

Segundo Chiyoshi *et al.* (2000), o interesse por problema de localização cresceu de forma significativa a partir da década de 60, quando os primeiros modelos matemáticos começaram a ser desenvolvidos. O autor descreve que os modelos de localização eram caracterizados pela otimização de uma medida de utilidade, sujeita as restrições operacionais relevantes ao problema. Logo, o que se deseja é encontrar uma solução, se possível ótima, que minimize o custo total de instalações e transportes (BALLOU, 2006).

Os problemas de localização de instalações podem ser categorizados mediante ao seu objetivo e de acordo com as características do espaço de soluções. Eiselt e Laporte (1995) define a existência de três tipos de objetivos. São eles: (1) a minimização em função da distância, atrelados a serviços que são desejáveis e o quanto mais próximos dos clientes maior é o retorno da função, (2) maximização em função da distância, voltados a serviços indesejáveis e quanto mais longe dos clientes mais é satisfatório, e por fim, (3) o balanceamento em função da distância, visa encontrar o equilíbrio entre a disposição das instalações com relação a seus clientes.

Nesse sentido, Fernandes *et al.* (2017) afirmam que, com base nesses objetivos, é possível agrupar os problemas de localização em um plano Euclidiano ou em rede. No qual, o autor descreve que os modelos de localização no plano, comumente, envolvem problemas em que as possibilidades de localização de instalações são infinitas. Já os problemas em rede são contínuos ou discretos. Contínuos, se a localização das instalações se encontra nos vértices ou em qualquer ponto dos arcos. Discretos, se a localização das instalações só pode ser realizada unicamente nos vértices.

Segundo Guazzelli e Cunha (2015) o modelo mais utilizado para solucionar problemas de localização, são os discretos, em que consistem na estruturação de conjuntos dados pré-

selecionados dos possíveis locais que deverão ser utilizados para instalar as unidades. Church e Revelle (1974) corroboram com essa afirmação, definindo que os possíveis pontos de alocação de uma instalação, não existem em qualquer lugar do espaço, geralmente são levantados através de estudos e análises prévias.

Contudo todas essas metodologias de alocar e classificar o problema de localização de instalações acaba se tornando mais complexo quando é adicionado a ideia de instalações públicas e privadas. Cada grupo de instalações tem seus critérios a serem atendidos e alguns deles acabam se tornando mais relevantes para um grupo do que para outro.

2.2.1 Diferenças na localização de instalações públicas e privadas

Os estudos voltados a localização das instalações públicas a cada dia vêm recebendo mais atenção e foi estabelecido como um campo de estudo tão importante quanto as análises realizadas na localização das instalações privadas (GUNAWARDANE, 1982; MURRAY, 2016). Segundo Mapa e Lima (2012) os problemas de localização de instalações ocorrem tanto no setor público como no privado e em uma visão macro, essas diferenças não são tão claras.

Nesse sentido, o objetivo macro do problema de localização consiste em alocar a instalação de forma a obter a menor distância até o ponto de demanda para conseguir minimizar os custos de transporte e maximizar a área de cobertura (MAPA E LIMA, 2012). Obtendo assim, uma maior acessibilidade para instalação e reduzindo os custos mediante a quantidade de unidades que serão instaladas em uma determinada região. Nesse sentido, a localização das unidades públicas e privadas resultam em um elemento chave para o seu planejamento estratégico e execução de suas atividades (OWEN e DASKIN, 1998).

Segundo Revelle *et al.* (1970) em uma visão micro, essas diferenças acabam se tornando mais claras no momento da definição dos objetivos e restrições que são formulados para o setor público e privado, influenciando na concepção do ótimo para cada setor. Church e Revelle (1974) apresentam que na localização de instalações privadas, um dos principais objetivos é o de minimizar os custos de fabricação e distribuição de seus bens e serviços. Porém, o autor destaca que ao contrário das localizações privadas, os objetivos da análise de localização de instalações públicas são mais difíceis de serem enxergados e quantificados.

Essa dificuldade consiste, pois, as decisões tomadas pelo serviço público são construídas por um conjunto de diferentes proprietários, voltados para a sociedade como um todo e o objetivo do setor público é maximizar um benefício ou minimizar um custo que não é quantificável em termos de dinheiro (REVELLE *et al.*, 1970). Logo, a estratégia proeminente

do setor público, é a seleção de um conjunto ideal de locais de um conjunto de candidatos, que atendam a toda população ou a máxima população possível (GUNAWARDANE, 1982).

Tendo em vista essa dificuldade na quantificação dos dados para a modelagem do problema de localização de instalações públicas, foram definidas duas medidas que possibilita uma maior visualização do problema para o tomador de decisão. A primeira delas, é voltada à distância total ponderada ou tempo para viajar até as instalações, e a segunda, é voltada a distância ou o tempo que o usuário mais distante de uma instalação teria que percorrer para chegar a essa instalação, em outras palavras seria a distância máxima de serviço (CHURCH e REVELLE, 1974).

Chiyoshi *et al.* (2000) na busca por melhor definir o problema de localização de instalações públicas, definem que o problema pode ser classificado em duas categorias: localização de serviços não emergenciais e localização de serviços de emergência. Na primeira categoria o autor define que as instalações seriam aquelas que não tem urgência para o atendimento, como as agências de correio, serviços relacionados ao meio ambiente, escolas ou mesmo edifícios públicos. Já para a segunda categoria, o autor define que as instalações são as que realmente necessitam de uma atenção a mais e que seu atendimento tem um critério de urgência, como a localização de hospitais, serviços de atendimento a incidentes, estações do corpo de bombeiros e até mesmo de instalações policiais que compõem o quadro de segurança pública de uma região.

2.2.2 Localização de instalações voltadas a segurança pública

Quando se trata de localizações de instalações voltadas à segurança pública, um dos pontos mais importantes e de grande levantamento de discussão é a localização de instalações policiais para o combate ao crime e a prevenção da ordem. Mediante as classificações supracitadas para problemas de localização, esse problema se encaixa em serviços públicos emergenciais, que possuem como objetivo comum, o atendimento da população em um menor espaço de tempo possível, já que visam salvar vidas e coibir crimes (SILVA *et al.*, 2013).

Nesse sentido, Chiyoshi *et al.* (2000) corroboram que uma das medidas mais utilizadas para solucionar o problema, é a máxima distância a ser percorrida (chamada de distância crítica) entre qualquer usuário do sistema e a instalação (ou servidor) mais próxima. Logo, o desempenho das unidades policiais está atrelado ao seu posicionamento com relação aos nós de demanda em uma rede. Em outras palavras, o objetivo central desse tipo de problema é alocar a instalação policial em um ponto que minimize o tempo médio de resposta entre o recebimento

de uma ligação e a chegada do veículo de emergência no local da demanda em questão (BANDYOPADHYAY *et al.*, 2012; GUSMÃO *et al.*, 2019).

Além desse fato, Gurgel *et al.* (2010) complementam que a alocação das instalações policiais deve ser baseada próximo a pontos em que existe um grande número de ocorrências, dessa forma, devem ajudar na redução da criminalidade e aumentar a cobertura policial na região. Mediante a esse cenário, é possível aumentar a cobertura reduzindo os custos na construção de novas unidades, ou seja, concentrando a maior parte das instalações em locais que detém de uma maior demanda e evitando gastos na criação de instalações em locais que contém pouca ou nenhuma demanda (GUSMÃO *et al.*, 2019).

Realizando esse mapeamento das áreas que contém maior demanda e as que tem menos demanda, é possível criar as divisões geográficas para que as instalações policiais definam até onde será sua área de patrulha. Segundo Curtin *et al.* (2007) essas divisões geográficas ótimas, podem alocar de forma eficiente a mão de obra policial limitada e outros recursos, elevando a probabilidade de êxito das instalações para o combate ao crime.

Porém, encontrar essas divisões geográficas ótimas que maximizem a área de cobertura das instalações policiais nem sempre é fácil, pois, dificilmente se utilizam métodos formais com uma base metodológica. Uma vez que, as fronteiras geográficas que limitam a área de atuação da polícia, são historicamente desenhadas a mão com base em conhecimentos a priori de um oficial ou gestor (MITCHELL, 1972; TAYLOR e HUXLEY, 1989).

É necessário a criação de modelos matemáticos que levem em consideração os pontos de demanda (crimes ou mesmo ocorrências) ao longo de uma região e as limitações dos recursos (instalações policiais e viaturas) existentes. Caso o contrário, essa falta de procedimentos formais e de medidas quantitativas irá tornar mais difícil a tomada de decisão da alta gerência policial, além de poder conduzir a escolhas menos eficientes do ponto de vista econômico e operacional (TAYLOR; HUXLEY, 1989).

2.3 Modelos voltados a localização de instalações

Os modelos de localização de instalações hoje, estão bem estruturados e difundidos nas mais diversas áreas da engenharia, com amplas aplicações e mais de um século de história na academia científica (FARAHANI *et al.*, 2019). Porém, esse cenário só foi alcançado mediante a complexidade dos problemas de localização de instalações, que nos últimos anos cresceram, levando à necessidade de desenvolvimento de modelos matemáticos que gerassem soluções ótimas (GURGEL, 2010; MITCHELL, 1972; CORDEAU *et al.*, 2018; MURRAY, 2016).

Nesse sentido, visto que os modelos de localização buscam encontrar os melhores locais candidatos para alocar as instalações mediante a localização dos clientes (KCHAOU BOUJELBEN e BOULAKSIL, 2018), foi despertado um grande interesse por parte dos pesquisadores que criaram uma vasta literatura tratando dessa problemática. Os profissionais de pesquisa operacional desenvolveram diversos modelos de programação matemática para solucionar os mais diferentes tipos de problemas de localização (OWEN e DASKIN, 1998).

Owen e Daskin (1998) apresentam uma revisão da literatura sobre alguns dos principais modelos voltados ao problema de localização de instalações estáticas e determinísticas em torno das diferentes funções objetivas requeridas por aplicações comuns. Classificando-os de acordo com os tipos de problemas mais recorrentes, o autor destaca os seguintes modelos: *P-median problem*, *Center problems* e o *Covering problems*.

2.3.1 P-median problem

O *P-median problem* é um modelo voltado para solucionar o problema de localizar p instalações em uma rede, em que a distância média percorrida a partir dos pontos de demanda até as instalações de serviço mais próxima é minimizada (NEEMA *et al.*, 2011). A sua estrutura matemática segundo Owen e Daskin (1998), pode ser apresentada pelos seguintes conjuntos de equações que vão da 2.1 a 2.6.

$$\text{Minimizar } \sum_i \sum_j h_i d_{ij} Y_{ij} \quad (2.1)$$

$$\text{Sujeito à: } \sum_j X_j = P, \quad (2.2)$$

$$\sum_j Y_{ij} = 1, \quad (2.3)$$

$$Y_{ij} - X_j \leq 0, \quad \forall i, j, \quad (2.4)$$

$$X_j \in \{0,1\} \quad \forall j, \quad (2.5)$$

$$Y_{ij} \in \{0,1\} \quad \forall i, j. \quad (2.6)$$

Onde os parâmetros são definidos como:

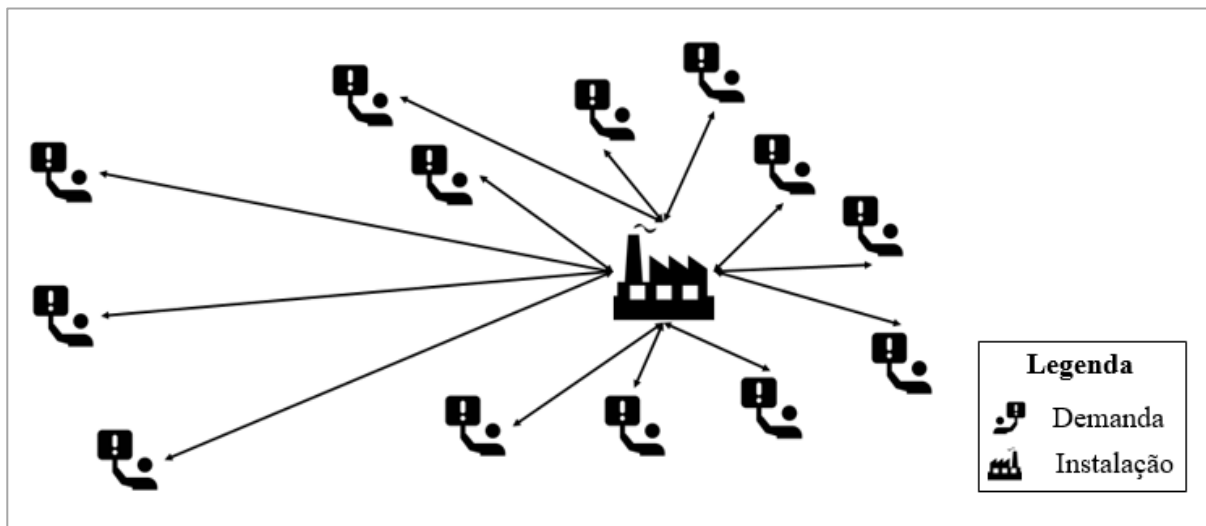
- ✓ i = índice do nó de demanda
- ✓ j = índice do local da potencial instalação
- ✓ h_i = demanda no nó i
- ✓ d_{ij} = distância entre o nó de demanda i e o local da instalação potencial j

- ✓ P = número de instalações a serem localizadas
- ✓ $X_j = \begin{cases} 1, & \text{se localizarmos no local da instalação potencial } j; \\ 0, & \text{se não.} \end{cases}$
- ✓ $Y_{ij} = \begin{cases} 1, & \text{se as demandas no nó } i \text{ forem atendidas por um recurso no nó } j; \\ 0, & \text{se não.} \end{cases}$

Nesse sentido, a Equação (2.1) define o objetivo do modelo que é, minimizar a distância total ponderada pela demanda entre clientes e instalações. A Equação (2.2) exige que exatamente as instalações P estejam localizadas. Já a Equação (2.3) garante que toda demanda seja atribuída a algum local com instalação. No qual, a Equação (2.4) permite apenas a atribuição, em locais nos quais as instalações podem ser localizadas. Por fim, as Equações (2.5) e (2.6) são requisitos binários para as variáveis do problema.

Em outras palavras, o modelo busca encontrar a melhor localização das instalações de modo a minimizar a distância média percorrida e ao mesmo tempo que atenda toda a demanda. Esse objetivo pode ser ilustrado na Figura 2.

Figura 2 – Ilustração de localização pelo P -median problem



Fonte: O Autor (2019)

O primeiro artigo publicado voltado a essa temática foi elaborado por Hakimi. E é apresentado que o P -median problem permite trabalhar com múltiplos pontos de demanda e instalações, trabalhando para que os pontos de demanda sejam atendidos por no mínimo uma das instalações (HAKIMI, 1964). Segundo Aguado e Trandafir (2012) esse é um dos modelos mais trabalhados na literatura quando se trata de otimização combinatória, em que existem muitas aplicações e entre elas está a análise de problemas de localização. Segundo Stefanello *et al.* (2015) o P -median problem pode ser utilizado na localização de diversas instalações, como hospitais, escolas e coleta de lixo, entre outros.

Para solucionar esse problema, diversos métodos foram desenvolvidos e a maioria dos algoritmos tradicionais é baseada em programação matemática (NEEMA *et al.*, 2011), utilizando desde programações lineares (REVELLE E SWAIN, 1970) a heurísticas e meta-heurísticas (ALCARAZ *et al.*, 2012). É possível verificar mais detalhes em uma extensa bibliografia sobre métodos de solução para o modelo *P-median problem* apresentado por Reese (2006).

Logo, cada problema de localização tem um objetivo específico e em alguns casos se faz necessário a busca por novos modelos que se adequem a sua problemática. Como é o caso dos problemas de centro (*Center problems*).

2.3.2 Center problems

Os *Center problems* são modelos de grande importância na área de ciência da computação teórica e possui diversas aplicações em vários campos de estudo. E um dos campos de estudos no qual o modelo vem a se enquadrar, são os problemas de localização de instalações, que foram amplamente trabalhados nas últimas décadas (DING *et al.*, 2017).

Os problemas de localizações podem ser apresentados em termos de escolhas de centros ou instalações para melhorar o atendimento de um determinado conjunto de clientes, no qual, se deseja minimizar o máximo possível a distância média de um cliente ao seu centro atribuído (NA *et al.*, 2014). Esse é o caso dos modelos de *Center problems*, que também são chamados de problemas MiniMax e seu principal objetivo é minimizar a máxima distância entre os pontos de demanda e a instalação mais próxima (DAVOODI, 2019; DASKIN, 2001; FARAHANI *et al.*, 2010). A sua estrutura matemática é semelhante ao do modelo *p-median problems*, que pode ser descrito através das seguintes equações que vão da 2.7 a 2.13.

$$\text{Minimizar } D \quad (2.7)$$

$$\text{Sujeito à:} \quad \sum_j X_j = P, \quad (2.8)$$

$$\sum_j Y_{ij} = 1 \quad \forall i, \quad (2.9)$$

$$Y_{ij} - X_j \leq 0, \quad \forall i, j, \quad (2.10)$$

$$D \geq \sum_j d_{ij} Y_{ij} \quad \forall i, \quad (2.11)$$

$$X_j \in \{0,1\} \quad \forall j, \quad (2.12)$$

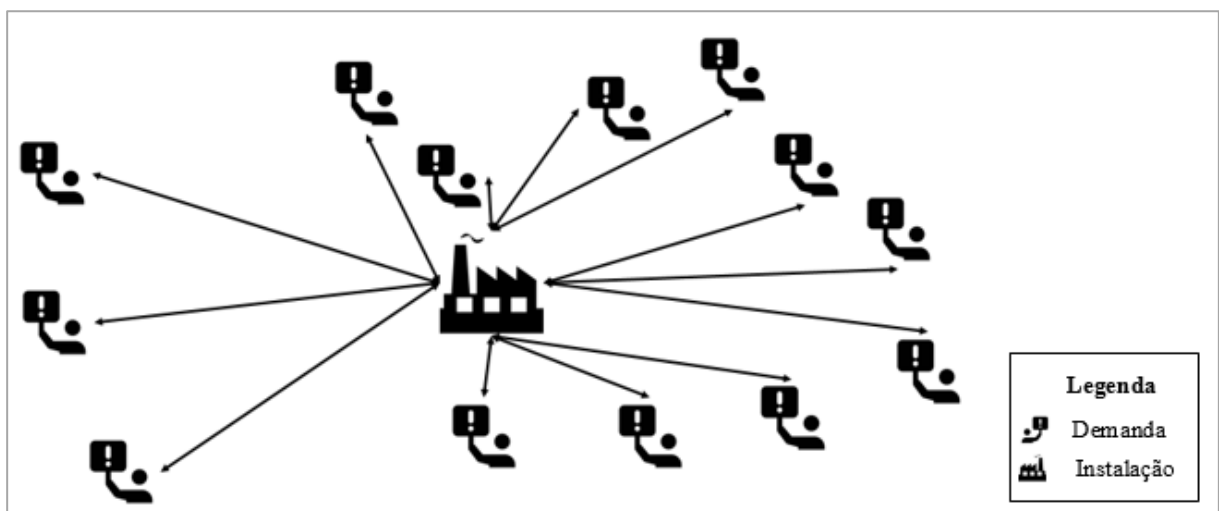
$$Y_{ij} \in \{0,1\} \quad \forall i, j. \quad (2.13)$$

Onde os parâmetros são definidos como:

- ✓ i = índice do nó de demanda
- ✓ j = índice do local da potencial instalação
- ✓ d_{ij} = distância entre o nó de demanda i e o local da instalação potencial j
- ✓ P = número de instalações a serem localizadas
- ✓ D = distância máxima entre um nó de demanda e a instalação mais próxima.
- ✓ $X_j = \begin{cases} 1, & \text{se localizarmos no local da instalação potencial } j; \\ 0, & \text{se não.} \end{cases}$
- ✓ $Y_{ij} = \begin{cases} 1, & \text{se as demandas no nó } i \text{ forem atendidas por um recurso no nó } j; \\ 0, & \text{se não.} \end{cases}$

Nesse sentido, na Equação (2.7) define o objetivo do modelo *center problems*, que é a minimização da máxima distância entre qualquer nó de demanda e sua instalação mais próxima. No qual, a semelhança entre o modelo *p-median* e o *center problems* está nas Equações (2.8), (2.9) e (2.10) que são idênticas as Equações (2.2), (2.3) e (2.4) do modelo *p-median*. Seguido para a equação (2.11) é apresentado a definição para se encontrar a máxima distância entre qualquer nó de demanda i e a instalação mais próxima j . Por fim, nas Equações (2.12) e (2.13) são requisitos binários para as variáveis do problema. Na Figura 3 é ilustrado o objetivo do modelo *Center problem*, que claramente busca centralizar a instalação.

Figura 3 – Ilustração de localização pelo Center Problems



Fonte: O Autor (2019)

Ding *et al.* (2017) apresentam que o modelo pode ser construído atribuindo aos vértices de entrada como os clientes e os centros selecionados como as instalações. Mapa e Lima (2012) complementam que quando as instalações estão restritas ao nó de uma rede, tem-se o problema

de centro de vértice. E se caso seja permitido que as instalações possam ser alocadas em qualquer lugar da rede, o problema é de centro absoluto.

Entretanto, não são todos os casos que o mais adequado, é o objetivo de minimização da distância média percorrida ou mesmo a minimização da máxima distância, como é o caso dos modelos de cobertura, chamados de *Covering problems*.

2.3.3 Covering problems

Os modelos de problema de cobertura (*covering problems*) englobam uma grande parte dos problemas voltados localização de instalações. Esses problemas incorporam uma noção de proximidade, ou seja, um raio de cobertura, que determina se um ponto da demanda será coberto ou não por uma instalação em potencial, e é dito que um ponto de demanda é coberto por uma instalação se mesmo estiver dentro do raio de cobertura dessa instalação (CORDEAU *et al.*, 2018).

Nesse sentido, é realizado uma análise de uma determinada região para definir a localização de um grupo de instalações que irão cobrir a demanda dessa região. É o caso das escolas públicas, delegacias, bibliotecas, hospitais, edifícios públicos, correios, parques, centros militares, instalações de radar, agências bancárias, centros comerciais e instalações de eliminação de resíduos, no qual, podem ser formulados como *covering problems* (FRANCIS e WHITE, 1974; FARAHANI *et al.*, 2012; GARCÍA e MARÍN, 2015; SNYDER, 2011, MURRAY, 2016).

Nos modelos de *Covering problems*, um ou mais clientes irão caracterizar a demanda de uma região, em que os mesmos devem ser atendidos por pelo menos uma das instalações dentro de uma determinada distância crítica (FARAHANI *et al.*, 2012; GUAZZELLI e CUNHA, 2015). Esse problema está especialmente ligado aos serviços de instalações emergenciais (delegacias e hospitais) e não necessariamente precisam ser atendidos pela instalação mais próxima (CHIYOSHI *et al.*, 2000; GURGEL *et al.*, 2010; CHURCH e REVELLE, 1974).

Nesse sentido, a distância crítica está relacionada ao tempo que será gasto entre a instalação receber um chamado e encaminhar o serviço (ambulância ou viatura) para o local de demanda. É importante ressaltar que segundo Curtin *et al.* (2007), não existe uma única medida explícita que determine qual será essa distância ou tempo de resposta, em que essas instalações emergenciais venham a reagir de forma admissível, pois diversos fatores influenciam nessa questão. Mas, é notória a existência dessas diversas medidas subjetivas e objetivas na prestação

de um serviço desse tipo, e o tempo de resposta é um dos mais importantes (BROWN e COULTER, 1983).

Assim, mediante a essas dificuldades Schilling *et al.* (1993) definem que os modelos de *Covering problems* podem ser classificados em duas categorias: *Set Covering Problem* (SCP) e *Maximal Covering Location Problem* (MCLP).

2.3.3.1 Set Covering Problem

O modelo *Set Covering Problem* (SCP) tem como objetivo, a minimização dos custos associados a localização das instalações de forma que um determinado nível de cobertura seja atingido, e dado que o custo para a construção de cada instalação é o mesmo, o objetivo é equivalente a minimização do número de instalações a serem alocadas (OWEN e DASKIN, 1998). O objetivo do *Set Covering Problem* pode ser representado em sua estrutura matemática através das seguintes equações que vão da 2.14 a 2.16.

$$\text{Minimizar } \sum_j c_j X_j \quad (2.14)$$

$$\text{Sujeito à: } \sum_{j \in N_i} X_j \geq 1 \quad \forall i, \quad (2.15)$$

$$X_j \in \{0,1\} \quad \forall j, \quad (2.16)$$

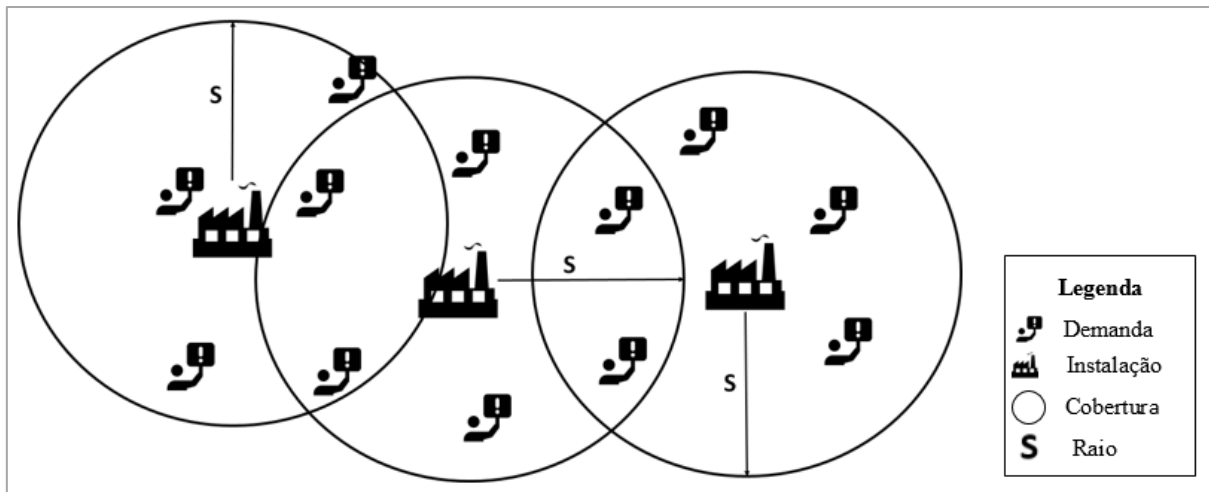
Onde os parâmetros são definidos como:

- ✓ c_j = custo fixo de localização de uma instalação no nó j
- ✓ S = máxima distância de serviço aceitável (ou tempo)
- ✓ N_i = conjunto de locais de instalação j dentro da distância aceitável do nó (*i. e.*, $N_i = \{j | d_{ij} \leq S\}$)
- ✓ $X_j = \begin{cases} 1, & \text{se localizarmos no local da instalação potencial } j; \\ 0, & \text{se não.} \end{cases}$

Nesse sentido, na equação (2.14) temos o objetivo do modelo *set covering problem*, que se resume em minimiza o custo da localização da instalação ou mesmo, minimizar o número de instalações a serem localizadas. E na equação (2.15), é exigido que todos os pontos de demanda tenham pelo menos uma instalação localizada dentro da distância de serviço aceitável. Por fim, a equação (2.16) exigem integralidade para as variáveis de decisão. Na Figura 4 é possível ilustrar o objetivo do SCP, no qual, o nível de cobertura para esse exemplo é de 100% da demanda.

Contudo, na maior parte dos casos o que acontece é que, o decisor percebe que seus recursos não são suficientes para atingir um nível de cobertura que atenda a todos os clientes, dessa forma, o tomador de decisão pode procurar cobrir o maior número possível de clientes dentro de sua área de cobertura, utilizando de seus recursos limitados (MURRAY, 2016). Nesse sentido, foi formulado por Church e Reville em meados da década de 1970 o modelo *Maximal Covering Location Problem* (MCLP) utilizando dos recursos da Pesquisa Operacional em programação linear e inteira (CORDEAU *et al.*, 2018).

Figura 4 – Ilustração de localização pelo Set Covering Problem



Fonte: O Autor (2019)

2.3.3.2 Maximum Covering Location problem

O modelo *Maximum Covering Location problem* (MCLP), vem como uma solução para os problemas de localização de instalações que necessitam de cobrir uma distância de serviço desejada pelo tomador de decisão. Porém, o tomador de decisão não pode apenas se concentrar em cobrir uma região que atenda o maior número de clientes, mas também com o tempo de resposta para a prestação do serviço oferecido para os clientes dentro dessa região de cobertura (CHURCH e REVELLE, 1974).

Nesse sentido, a distância de serviço na maior parte dos casos está atrelada ao tempo de resposta entre o deslocamento da instalação até o ponto de demanda (CORDEAU *et al.*, 2018). Significa que, não adianta estabelecer uma distância S muito elevada e que venha a cobrir mais pontos de demandas, dado que, o tempo de resposta para chegar no ponto mais distante inviabiliza a qualidade do serviço prestado na região, como é o caso de delegacias, hospitais e corpos de bombeiros (MURRAY, 2016; CHURCH e REVELLE, 1974; OWEN e DASKIN, 1998).

Segundo Murray (2016) o modelo basicamente funciona como uma rede de nós e arcos que representam o espaço geográfico do estudo, no qual, os nós são as concentrações de demanda pelo serviço da instalação e os arcos refletem a distância entre os nós. Dado que o modelo seja aplicado em um plano, ou seja, uma região territorial, a cobertura de uma instalação usando a distância euclidiana é representada por um círculo de raio de S contando com a instalação no centro (BANSAL e KIANFAR, 2017).

Assim, um problema que contenha múltiplos nós de demanda concomitantemente dispersos no espaço, pode acarretar uma solução que detenham de uma grande quantidade de instalações, o que torna o projeto inviável do ponto de vista orçamentário (GUSMÃO *et al.*, 2019). Logo, o MCLP pode ser aplicado para localizar um número fixo de instalações na rede de forma a maximizar o número de nós de demanda que são cobertos dentro de uma de distância ou tempo especificado pelo tomador de decisão (SNYDER e HAIGHT, 2016).

Nesse sentido, ao realizar uma pesquisa mais técnica e aprofundada sobre o problema de localização de máxima cobertura. É apresentado, a estrutura matemática do modelo MCLP que foi proposto por Church e ReVelle em 1974 através das seguintes equações que vão da 2.17 a 2.21.

$$\text{Maximizar } \sum_{i \in I}^n a_i y_i \quad (2.17)$$

$$\text{Sujeito à: } \sum_{j \in N_i} x_j \geq y_i \quad \forall i \in I \quad (2.18)$$

$$\sum_{j \in J} x_j = P \quad (2.19)$$

$$x_j \in \{0,1\} \quad \forall j \in J \quad (2.20)$$

$$y_{ij} \in \{0,1\} \quad \forall i \in I \quad (2.21)$$

Onde os parâmetros são definidos como:

- ✓ I = indica o conjunto de nós de demanda;
- ✓ J = indica o conjunto de potenciais locais de operação;
- ✓ $N_i = \{j \in J \mid d_{ij} \leq S\}$;
- ✓ d_{ij} = menor distância entre o nó i e o nó j ;
- ✓ S = valor máximo estabelecido da distância entre o nó de demanda e o nó da operação (distância de serviço desejada);

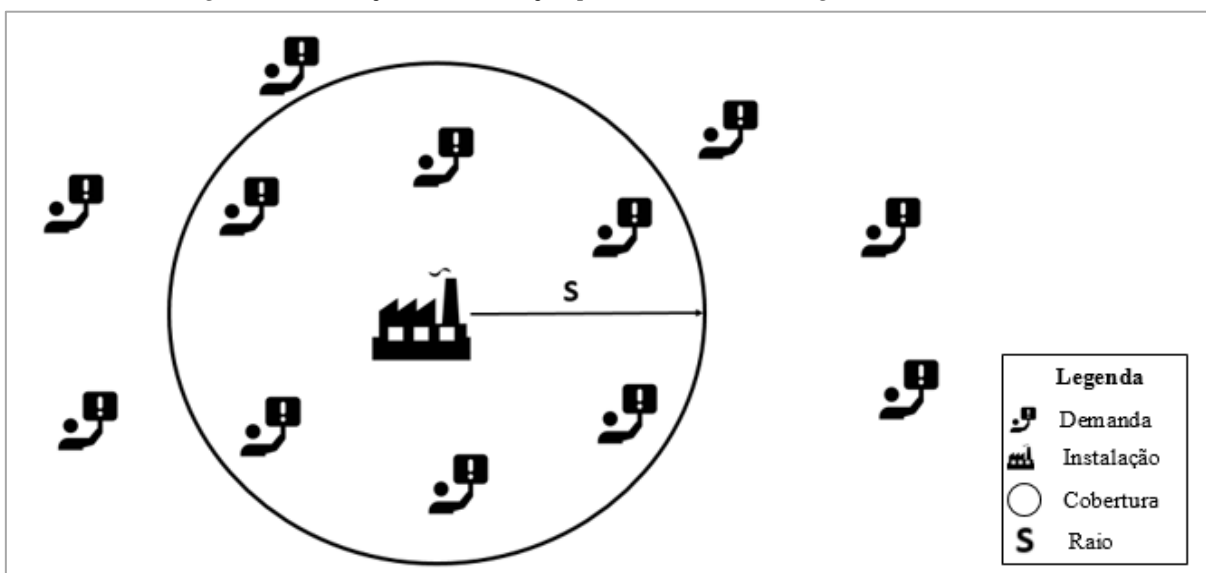
- ✓ a_i = população atendida no nó de demanda i ;
- ✓ P = número de operações a serem alocadas.
- ✓ $x_j = \begin{cases} 1, & \text{se uma operação está alocada no nó } j; \\ 0, & \text{caso contrário.} \end{cases}$
- ✓ $y_i = \begin{cases} 1, & \text{se um nó de demanda } i \text{ é coberto por pelo menos um nó de potencial } j; \\ 0, & \text{caso contrário.} \end{cases}$

N_i é o conjunto de locais candidatos para fornecer cobertura para um nó de demanda i . Um nó de demanda é chamado de "coberto" quando a instalação mais próxima desse nó está a uma distância menor ou igual a S . Nesse sentido, é chamado de "descoberto" quando a instalação mais próxima desse nó está a uma distância maior que S .

Analisado a função objetivo descrita na Equação (2.17), temos que o objetivo é maximizar o número de nós de demanda cobertos dentro da distância de serviço desejada. Nas equações seguintes são apresentadas as restrições do problema. A Equação (2.18) permite y ser igual à 1 somente quando uma ou mais operações forem estabelecidas em locais dentro do conjunto N (isto é, uma ou mais operações estão alocadas dentro da distância S em relação ao ponto de demanda i). Já na Equação (2.19) é definido o número de instalações alocadas no qual, está limitado por P , que é definido pelo decisor.

Por fim, nas Equações (2.20) e (2.21) é indicado que apenas valores inteiros podem fazer parte da solução, apresentando assim, um o modelo é de programação inteira e binária. Na Figura 5 é ilustrado o objetivo do MCLP, dado que a quantidade recursos disponíveis possibilita apenas a alocação de uma instalação.

Figura 5 – Ilustração de localização pelo Maximal Covering Location Problem



Fonte: O Autor (2019)

Nesse sentido, é possível investir os recursos de forma mais eficiente, alocando as instalações de modo a atingir a maior parte da demanda com os recursos disponíveis. O modelo depois de executado apresenta uma solução completa, especificando o valor máximo da demanda coberta pela melhor ou ótima localização das instalações (dependendo do método de solução do problema) e principalmente indicando quais são elas.

Segundo Church e ReVelle (1974) o problema do MCLP pode ser solucionado tanto por uma heurística como por uma programação linear. Nesse sentido, caso o problema seja solucionado por programação linear, o ótimo será encontrado. Contudo, se o problema for respondido por uma heurística, é possível encontrar um conjunto de melhores soluções, no qual, o ótimo pode ou não, ser encontrado dentro desse conjunto.

Segundo Cordeau *et al.* (2018) na maior parte dos casos, o número de pontos de demanda e locais candidatos é relativamente pequeno e podem ser resolvidas de forma satisfatória pelas heurísticas existentes ou por *solvers* de uso geral que são programados em diversas linguagens. Já em outros casos específicos, os autores retratam que o número de pontos de demanda pode chegar aos milhões. Corroborando com essa afirmação Máximo *et al.* (2017) discorrem que mesmo com os avanços tecnológicos, as aplicações em larga escala do MCLP em programação linear é um desafio, pois com o aumento de nós de demanda, vão surgindo algumas limitações com relação a capacidade de processamento em um curto espaço de tempo.

Pesando nessas dificuldades foram desenvolvidas algumas heurísticas que procuram reduzir o tempo de processamento dos dados nesses casos específicos. Logo, com algumas heurísticas atuais é possível solucionar um problema dessa magnitude em frações de segundos, onde seu valor objetivo é pelo menos 91% do objetivo ótimo em 90% dos casos (BANSAL e KIANFAR, 2017). Church e ReVelle (1974) apresentam as duas primeiras heurísticas voltadas a solução desse problema, o algoritmo *Greedy Adding* (GA) e o algoritmo *Greedy Adding with Substitution* (GAS), além dessas heurísticas os autores também demonstram as considerações que devem ser realizadas para trabalhar o problema de MCLP em programação linear para se atingir o ótimo.

Analisando o problema nas duas óticas através da literatura. É factível a evolução com relação a aplicação do modelo MCLP, isso, por se tratar de um modelo de programação linear inteira e binária, permitindo que uma gama de softwares privados e livres sejam usados para a solução desse problema (MURRAY, 2016; MÁXIMO *et al.*, 2017). Church e ReVelle (1974) realizam seus estudos na década de 70 com o Sistema de Programação Matemático (MPS) no

computador IBM 360 *Model 91* para obter soluções de programação linear ótimas para problemas de cobertura máxima.

Atualmente o IBM ILOG CPLEX *Optimization Studio* é um *software* bastante utilizado para solucionar problemas de MCLP pelos acadêmicos (GUSMÃO *et al.*, 2019), isso acontece pois o *software* é privado e sendo disponibilizado apenas para estudantes com algumas limitações (MÁXIMO *et al.*, 2017; FAZEL ZARANDI *et al.*, 2011). Além disso, o MCLP também foi integrado em vários pacotes de *software* comerciais baseados em sistema de informação geográfica (GIS), incluindo ArcGIS e TransCAD, dada sua utilidade geral (CURTIN *et al.*, 2007; CROMLEYA, 2012).

Nesse sentido, torna-se notória a existência de uma grande variedade de *softwares* e modelos que podem ser utilizados para solucionar os problemas de localização. Segundo Krivulin (2017) essa diversidade de modelos está relacionada as diferentes configurações que podem ser realizadas entre os modelos com base nos métodos de programação matemática e de otimização discreta.

2.3.4 Outros modelos abordados na literatura

Farahani *et al.* (2014) apresentam mais um modelo diferente, chamado de *hierarchical facility location problem* (HFLP) que tem por objetivo alocar instalações em diferentes níveis, de forma que as instalações em um nível superior possam atender eficazmente (tanto em objetivos financeiros como em acessibilidade) as instalações de níveis mais baixos. O autor ainda complementa que as instalações dentro dos diversos níveis não devem ser interpretadas como entidades isoladas e, portanto, a sua localização deve ser realizada mediante ao seu impacto no sistema.

Outro modelo é *Hub location problems* (HLP) que segundo Alumur e Kara (2008), estão relacionados com a localização de instalações *Hubs*, que são instalações especiais que servem como pontos de comutação, transbordo e classificação em sistemas de distribuição, em vez de servir diretamente a cada par origem-destino, as instalações *Hubs* concentram fluxos para aproveitar as economias de escala. Nesse sentido, o objetivo do modelo é a localização de instalações de *hub* e à alocação de nós de demanda a *hubs*, para rotear o tráfego entre pares de origem e destino, buscando minimizar os custos totais, maximizar a utilização dos transportadores e maximizar o nível de serviço (GELAREH e NICKEL, 2011; O'Kelly, 1987).

Também existem muitos modelos utilizados na literatura que são voltados a localizar instalações distantes da população, é o caso do modelo de Berman e Huang (2008). Os autores

apresentaram o modelo *Minimum Weighted Covering Location Problem With Distance Constraints* (MCLPDC) que tem o objetivo contrário ao *Maximal Covering Location Problem*. Nesse caso, o modelo visa minimizar os pontos de demanda serão cobertos pela área de atuação da instalação que será alocada, seguindo a região factível imposta pelas restrições com relação a distância mínima (Berman e Huang, 2008).

Porém, esses modelos voltados aos problemas de localizações de instalações necessitam de *inputs* para que sejam executados (HAKIMI, 1964; OWEN e DASKIN, 1998; CHURCH e REVELLE, 1974). *Inputs* que são, os pontos de demanda e os possíveis locais candidatos para a alocar uma instalação. Os pontos de demanda podem ser coletados em campo, mas os locais candidatos podem ser encontrados por diversos estudos (SILVA *et al.*, 2013, REVELLE e EISELT, 2005), como é o caso dos métodos de Clusterização (KAVEH *et al.*, 2010; OLIVEIRA, 2012).

2.4 Clusters como *inputs* para modelos de localização

A mineração de dados (*data mining*) é um termo utilizado na literatura para descrever a descoberta de conhecimento através de bancos de dados. Logo, a mineração de dados é um processo que pode fazer uso de técnicas como inteligência artificial, estatística, matemática e *machine learning*, para identificar e extrair informações úteis e subsequentes conhecimento em grandes bases de dados (TURBAN *et al.*, 2005).

Segundo Oliveira (2012), entre as ferramentas utilizadas pelo *data mining* para identificar e extrair informações uteis, está a classificação dos dados. Durante a mineração, os dados são analisados e classificados por um padrão ou característica em comum. Nesse sentido, para encontrar esses padrões são utilizados procedimentos de clusterização, que visam encontrar um conjunto de k pontos para atuar como centros e então atribuir cada ponto ao seu centro mais próximo, formando assim k *clusters* (KAVEH *et al.*, 2010; OLIVEIRA, 2012; FRIGGSTAD *et al.*, 2017).

Esse procedimento é fundamental nos modelos voltados ao problema de localização de n instalações, pois, pode ser sintetizado em localizar os k centroides, definindo os possíveis locais candidatos para posicionar as instalações (OLIVEIRA, 2012; SILVA *et al.*, 2013; RISSO *et al.*, 2015). Segundo Revelle e Eiselt (2005) esse pré-processamento para que sejam encontrados os locais candidatos é fundamental para a execução de um modelo, possibilitando encontrar uma solução que atenda a uma maior parte da demanda requerida pelo decisor.

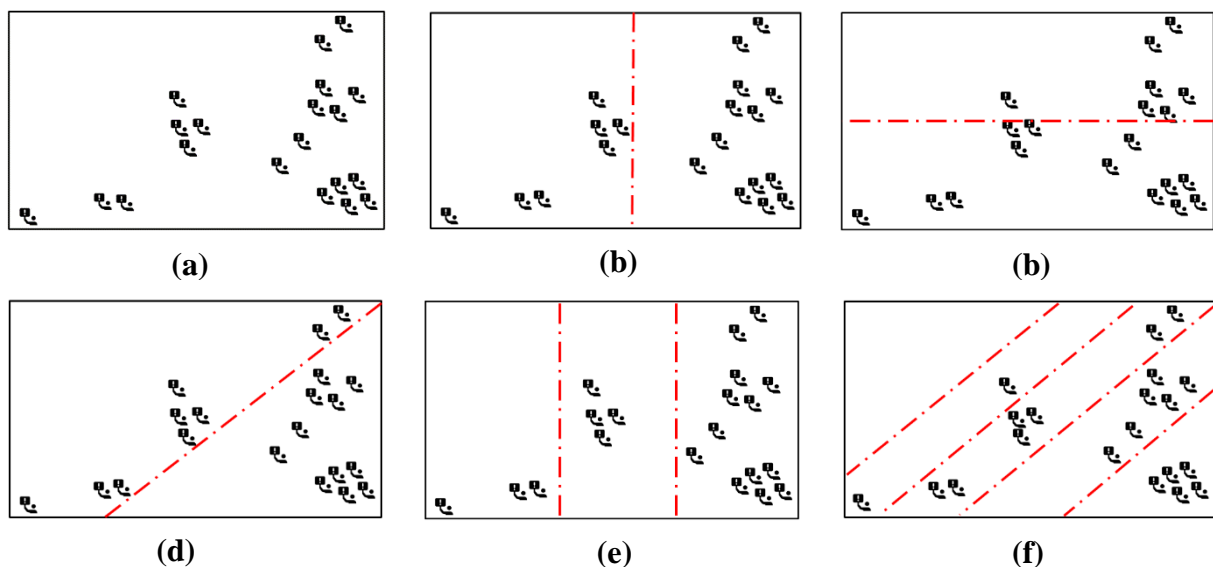
Nesse sentido, são identificados diversos métodos na literatura para realizar essa clusterização, mas o algoritmo *k-means* é o mais utilizado quando se deseja particionar n quantidades medidas em k *clusters* (MACQUEEN, 1967; HARTIGAN, 1975; PRAVEEN KUMAR *et al.*, 2019). Portanto, o algoritmo *k-means*, pode solucionar esse pré-processamento de forma relativamente rápida nos problemas de localização de instalações (OLIVEIRA, 2012; RISSO *et al.*, 2015).

2.4.1 Método *K-means*

Existem duas principais formas de se trabalhar com aprendizado de máquinas (*machine learning*), com o aprendizado supervisionado e não supervisionado (DAVID e SHWARTZ, 2014; LANTZ, 2015). Na aprendizagem supervisionada, é fornecido ao computador um conjunto de dados, onde o mesmo busca aprender para gerar um modelo que representa os relacionamentos e vínculos de dependência entre os dados de entrada e os resultados previstos na saída (PRAVEEN KUMAR *et al.*, 2019). Já no aprendizado não supervisionado, o computador classifica os dados com base em um conjunto de padrões e características, no qual, agrupamentos são formados para possibilitar a redução de dimensionalidade e detecção de anomalias entre os dados (PRAVEEN KUMAR *et al.*, 2019).

Nesse sentido, o algoritmo *k-means* que foi apresentado pela primeira vez por McQueen (1967) e é uma das técnicas de *machine learning* não supervisionada que conduz uma categorização exploratória dos dados até que as observações compartilhem de uma similaridade máxima no mesmo *cluster*, tendo a similaridade mínima entre os diferentes *clusters* (KOO *et al.*, 2019).

Figura 6 – Conjunto de demanda categorizado empiricamente em clusters



Fonte: O Autor (2019)

Nesse sentido, para facilitar o entendimento do procedimento de clusterização, na Figura 6a, é representado um conjunto de pontos de demanda que precisam ser categorizados em alguns *clusters*. Deste modo, essa categorização notoriamente pode ser analisada e entendida com diferentes cenários, através da Figura 6b, Figura 6c e Figura 6d, em que o conjunto de pontos de demanda foram categorizado em dois clusters através de uma linha vertical, horizontal e diagonal respectivamente. Podendo ainda, ser categorizado em mais *clusters*, como é apresentado na Figura 6e e Figura 6f, em três e quatro *clusters* respectivamente.

Porém esse procedimento de clusterização realizado na Figura 6 é realizado de forma empírica. Isso pode acarretar em não conseguir bons resultados, exigindo procedimentos formais para a clusterização, como é o caso do método *k-means* supracitado (LANTZ, 2015).

Segundo Silva *et al.* (2013) para uma melhor compreensão do método *k-means*, são realizadas algumas simplificações para apresentar o objetivo do método em sua estrutura matemática. No qual, pode ser apresentado com a função de custo (SILVA *et al.*, 2013; OLIVEIRA, 2012; HAIR *et al.*, 2005; DAVID e SHWARTZ, 2014) que contém todos os *clusters* e pode ser entendida pela Equação 2.22.

$$\text{custo} (kmeans) = \sum_{k=1}^k \sum_{x_i \in C_k} d(x_i, x_{0k}) \quad (2.22)$$

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} . Nesse sentido, o algoritmo *k-means* ao encontrar os *clusters* procura minimizar a soma das distâncias entre os pontos dados (ou pontos de demanda, caso seja um problemas de localização) ao centroide mais próximo (FRIGGSTAD *et al.*, 2017; MACQUEEN 1967; DAVID e SHWARTZ, 2014; OLIVEIRA, 2012; SILVA *et al.*, 2013; LANTZ, 2015; RISSO *et al.*, 2015). Consequentemente o método *k-means* consegue estruturar iterativamente um conjunto de *clusters* a partir de um conjunto de dados (KANUNGO *et al.*, 2002).

Segundo Silva *et al.*, (2013) o procedimento para realização da clusterização pelo método *k-means* segue os seguintes passos:

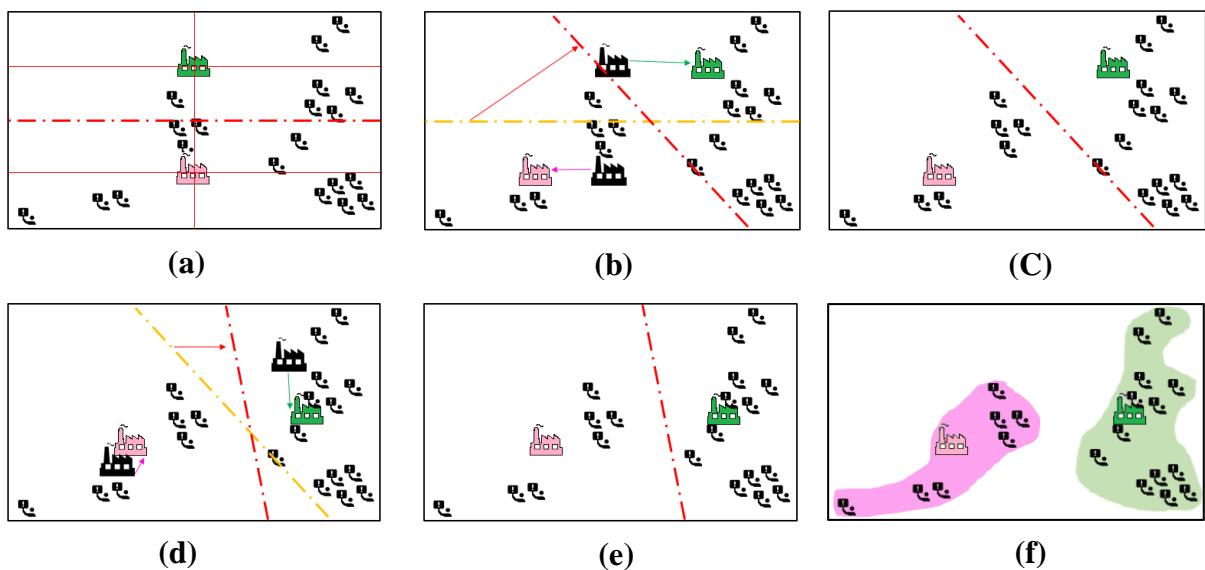
- Passo 1 - Selecionar k centroides aleatoriamente, como centroides iniciais de k *clusters* (*starting point*).
- Passo 2 - Atribuir cada ponto da demanda ao centroide mais próximo, baseado na distância euclidiana entre o objeto e o centroide. Depois atualizar os centroides, que

devem se mover para o centro dos clusters, de tal modo que k novos *clusters* serão criados.

- Passo 3 - Para cada k novos *clusters* criados, recalcular novos centroides para possibilitar nova atribuição dos pontos de demanda ao centroide mais próximo.
- Passo 4 - Repetir os passos 2 e 3, até que as atribuições dos centroides para seus respectivos *clusters* parem de mudar ou o número máximo de iterações seja atingido.

Esse procedimento iterativo pode ser mais claramente entendido com a Figura 7. No qual, contém um conjunto de imagens que representam os passos do método *k-means* para encontrar dois *clusters* e seus respectivos centroides com relação aos pontos de demanda.

Figura 7 – Ilustração dos passos do método *k-means*



Fonte: O Autor (2019)

Iniciando na Figura 7a, é executado o primeiro passo do método *k-means*, onde o conjunto de pontos de demanda é aleatoriamente dividido em dois *clusters*. Em seguida, o segundo passo é descrito na Figura 7b, no qual é atribuído para cada ponto da demanda o centroide mais próximo, forçando que os centroides sejam atualizados e se movam para o centro dos novos clusters como visto na Figura 7c. Chegando na Figura 7d, são recalculados os novos centroides para possibilitar uma nova atribuição dos pontos de demanda ao centroide mais próximo, gerando como resultado os novos *clusters* da Figura 7e. Por fim, na Figura 7f é possível visualizar os dois clusters encontrados pelo método *k-means* com seus respectivos centroides. Que em um problema de localização, os centroides podem ser descritos com locais candidatos para se atribuir uma nova instalação (SILVA *et al.*, 2013).

Corroborando com essa afirmação, Risso *et al.* (2015) aplicaram o método *k-means* com esse objetivo em uma planta industrial do segmento alimentício e verificaram que a abordagem é consistente, principalmente para os modelos de *covering problem*. E como o *k-means* pode ser relativamente escalável e eficiente para grandes conjuntos de dados, o método frequentemente termina encontrando um centroide ótimo, que minimizar a soma das distâncias entre (as instalações e os pontos de demanda) os pontos de dados (BERKHIN, 2002).

E é através desse procedimento formal do *k-means*, que a importância da clusterização de dados é amplamente reconhecida por diversos pesquisadores. Onde, é compreendido como um modelo para melhor gerenciar, navegar, filtrar e resumir de maneira eficiente e inteligente uma grande quantidade de dados, que possam servir para uma tomada de decisão com maior segurança (OYELADE *et al.*, 2010; FRIGGSTAD *et al.*, 2019; COHEN-ADDAD *et al.*, 2019; DHALMAHAPATRA *et al.*, 2019).

Mediante ao exposto, o *k-means* pode ser adicionado como ferramenta computacional de clusterização em sistemas de informações gerenciais que demandam de tratar grandes volumes de dados. Permitindo que os sistemas de informações gerenciais sejam executados com mais eficiência para ajudar os decisores em suas atividades do dia a dia. Nesse sentido, na próxima seção serão apresentados alguns dos conceitos do sistema de informação gerencial.

2.5 Sistema de informação gerencial

Os executivos e diretores organizacionais, quando estão analisando seus planos para localizar uma nova instalação, necessitam de gerenciar uma grande quantidade de dados e informações para tomar a decisão final de onde será alocada a instalação. Logo, a tomada de decisão pode ser entendida como o procedimento de seleção de algumas, ou apenas uma das alternativas factíveis (OLIVEIRA, 2004). Nesse sentido, a eficiência dessa tomada de decisão consiste na escolha da alternativa que, tanto quanto possível, ofereça os melhores resultados. (CARDOSO *et al.*, 2009).

Corroborando com essa afirmação, Jung (2004) ressalva que, com as melhores decisões é possível obter ações produtivas, rapidez na resolução dos problemas e melhores desempenhos. Porém, diversos fatores influenciam na tomada de decisão e entre eles estão as operações complexas que podem aumentar os custos dos erros, as rápidas mudanças no ambiente que produzem maiores incertezas, as crescentes regulamentações governamentais associadas à desestabilização política e as novas tecnologias que podem influenciar na distribuição da informação (TURBAN *et al.*, 2005).

Nesse sentido, para que as decisões organizacionais contornem esses fatores e sejam tomadas com rapidez e qualidade, se faz necessário que as organizações disponham de um sistema de comunicação eficiente, no qual, possibilite uma rápida disseminação da informação e do conhecimento por toda sua estrutura hierárquica (ANGELONI, 2003). Em síntese, um sistema de informação gerencial (SIG) pode ser a solução para essa problemática (STAIR e REYNOLDS, 2015).

Mediante a essa necessidade de aumentar a eficiência e eficácia das organizações, os SIGs começaram a ser desenvolvidos por volta dos anos de 1960. Onde são utilizados pelos gerentes para tomar decisões de problemas estruturados, através de um conjunto organizado de pessoas, métodos, *softwares*, banco de dados e *hardwares* (STAIR e REYNOLDS, 2015).

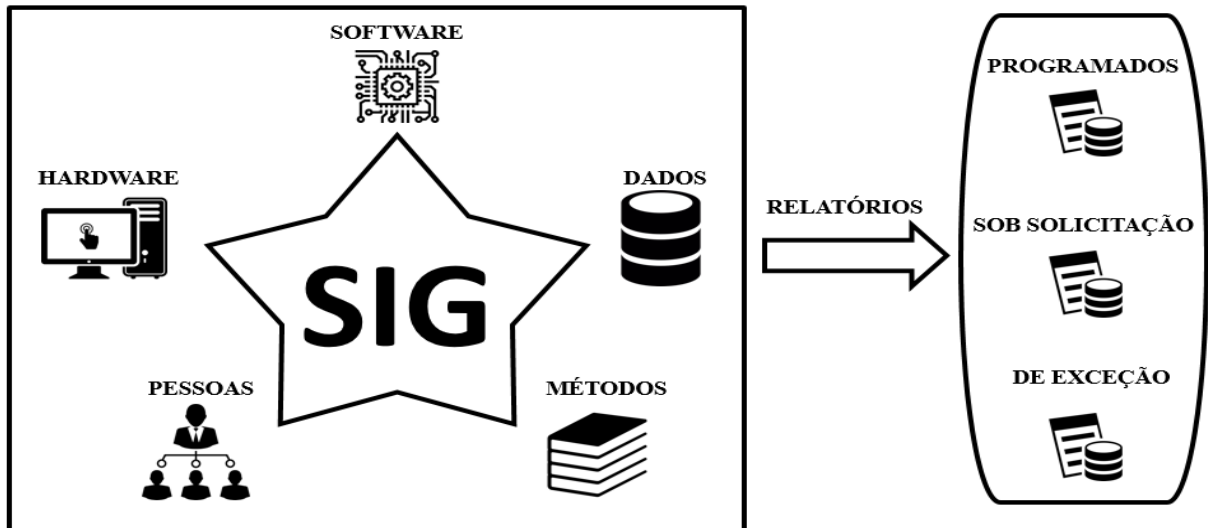
Em outras palavras, os sistemas de informações gerenciais facilitam o processo de captação, entrada e transformação dos dados em informações com valor agregado, para que possam ser utilizadas, com efetividade, nas tomadas de decisão (PORTO e BANDEIRA, 2006; BORGES e FERNANDES, 2016). Campeão *et al.* (2007) complementam que os SIGs fornecem através de seus *outputs*, informações sobre as operações internas e sobre o ambiente da empresa, visando assegurar o sucesso das estratégias organizacionais.

Nesse sentido, os principais *outputs* fornecidos por um SIG consistem de três tipos de relatórios, os programados, os sob Solicitação e os de exceção (STAIR e REYNOLDS, 2015), como pode ser observado na Figura 8. No relatório programado, é fornecido ao decisor quais são as mudanças que ocorreram dentro da organização em um intervalo de tempo, que pode ser estipulado diariamente, semanalmente ou mensalmente. O relatório sob solicitação, é produzido quando o decisor necessita de uma resposta rápida para a sua problemática, e dessa forma o SIG gera essa informação no momento em que a mesma é requerida. Já o relatório de exceção, é produzido quando uma situação incomum acontece, ou seja, o SIG gera o relatório quando algum indicador está fora de controle para que o decisor intervenha na problemática antes do problema tomar grandes proporções.

Na maior parte dos casos esses relatórios são orientados quase que exclusivamente aos eventos internos, e não aos eventos ambientais ou externos (DALFOVO e DALFOVO, 2009). O que apoia, principalmente, as funções de planejamento, controle e decisão no nível gerencial (LAUDON e LAUDON, 2004; STAIR e REYNOLDS, 2015).

Segundo Mithas *et al.* (2011), os SIGs podem ser entendidos como um dos principais pilares para que as organizações obtenham os melhores resultados. E, portanto, necessitam de ser bem implementados e gerenciados por toda organização (LI *et al.*, 2012; CANIËLS e BAKENS, 2012; MONACO e REDMON, 2012; MAKINDE *et al.*, 2001).

Figura 8 – Arquitetura básica do Sistema de Informação Gerencial



Fonte: O Autor (2019)

Corroborando com essa afirmação, Ferreira *et al.* (2015) complementam que o SIG é uma ferramenta indispensável para a obtenção de informações relevantes que podem trazer grandes benefícios para a organização. Propiciando relatórios mais precisos; mais rápidos; melhoria na produtividade; melhorias nos serviços realizados e oferecidos; maior agilidade na tomada de decisões; redução de custos nas operações e melhor acesso às informações (BORGES e FERNANDES, 2016)

Segundo Ploskas *et al.* (2015), essa arquitetura dos SIGs tem implicações gerenciais importantes, as quais permitem que os tomadores de decisão formulem seus estudos e façam, análises completas sobre o ambiente e a possível localização da nova instalação. Nesse sentido, um SIG possibilita um avanço na tomada de decisão em problemas de localização de instalações, garantido velocidade, confiança e segurança para os tomadores de decisão (GUSMÃO *et al.*, 2019).

Então, com base nos conhecimentos supracitado para os modelos de localização de instalações e alto volume de informações e dados que são exigidas para a sua solução, é notório a necessidade de um SIG para o desenvolvimento desta pesquisa. E simultaneamente por se tratar de um problema de localização de instalações públicas e emergências, ficou claro que o modelo *Maximal Covering Location Problem* possui o perfil adequado exigido pelas necessidades desta pesquisa. Principalmente pelo modelo já ter sido implementando em outra pesquisa semelhante, realizada por Curtin *et al.* (2007) que conseguiram ótimos resultados com aplicação do MCLP. Nesse sentido o próximo capítulo visa explorar com mais detalhes o modelo MCLP e avaliar seu histórico de aplicações para justificar o seu uso atualmente e nesta pesquisa.

3 REVISÃO DA LITERATURA NO MODELO MCLP

Para aprofundar os conhecimentos no modelo MCLP, que se mostrou o mais adequado para este estudo. As próximas seções visam apresentar a revisão da literatura a respeito do modelo empregado neste trabalho.

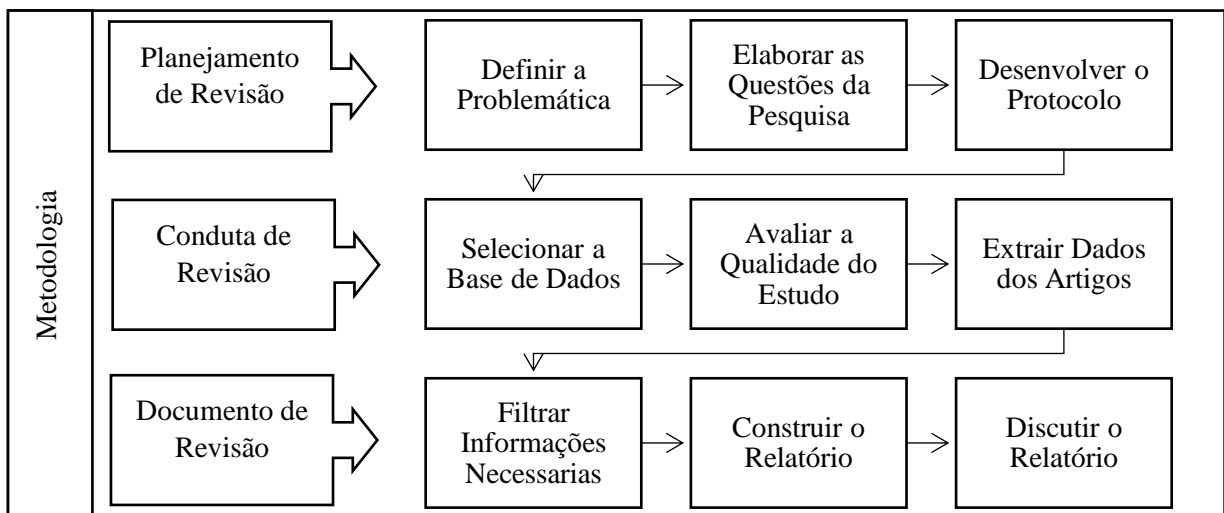
3.1 Método da Revisão

Para o desenvolvimento dessa seção e como forma de complemento do estudo desta pesquisa, foi realizado uma revisão da literatura, no qual, foi utilizado o método de revisão da literatura proposto por Brereton *et al.* (2007). Método esse, que é amplamente abordado na literatura quando se deseja realizar uma revisão de literatura mais estruturada. A proposta abordada por Brereton *et al.* (2007), sugere que a revisão seja elaborada com base em 3 fases: 1. Planejamento da Revisão, 2. Conduta a ser Seguida na Revisão, 3. Elaboração do Documento de Revisão.

É de extrema importância uma revisão da literatura para definir os limites de uma pesquisa, analisando sobre uma perspectiva científica e explorando todos os requisitos necessários como os autores, palavras chaves e suas respectivas bases de dados (DANE, 1990). Nesse sentido, a revisão da literatura busca analisar uma problemática mediante as referências publicadas em anais, jornais, periódicos e revistas, explorando as contribuições e a evolução da temática através do tempo (WEBSTER; WATSON, 2002).

Com base no modelo metodológico de Brereton *et al.* (2007), foi desenvolvido o seguinte modelo metodológico para essa pesquisa, apresentado na Fluxograma 1.

Fluxograma 1 – Processo de revisão da literatura



Fonte: O Autor (2019)

Nota: Adaptado Brereton *et al.* (2007)

3.1.1 Planejamento da revisão

Para que seja possível iniciar uma revisão da literatura, o escopo da problemática da pesquisa deve ser bem definido para assim guiar os primeiros passos. Nesse sentido, o principal objetivo dessa revisão é analisar os artigos publicados, que utilizem do modelo *Maximum Covering Location problem* (MCLP) para os mais diversos tipos de aplicações para o problema de localização, desde a sua criação até os dias atuais.

Seguindo o fluxo da metodologia, se faz necessário o levantamento das questões que vão nortear a revisão da literatura. Essas questões, serão respondidas durante a análise e elaboração do relatório de revisão na terceira fase. As respostas dessas questões visam servir de base para a justificativa do uso do modelo MCLP nesta pesquisa.

Questões:

1. Quantidade de artigos publicados sobre MCLP desde sua criação?
2. Como se comportou as publicações de MCLP nos últimos anos?
3. Quais são as principais fontes para se encontrar artigos de MCLP?
4. Quem são os autores mais relevantes sobre MCLP?
5. Como os autores mais relevantes de MCLP se comportaram ao longo do tempo com relação as publicações na área?
6. Quais são os países que estão à frente nas pesquisas de MCLP?
7. Quais artigos de MCLP tem maior relevância e foram mais citados ao longo do tempo?
8. Quais as diferentes aplicabilidades foram encontradas para o modelo de MCLP?

Para finalizar a primeira fase da metodologia, foi definido o protocolo da pesquisa visando a minimizar qualquer viés no estudo. Nesse sentido, temos a definição prévia do conteúdo da pesquisa e a garantia que não seriam processados artigos duplicados e artigos irrelevantes para a construção do documento de revisão voltado ao conteúdo da pesquisa.

3.1.2 Conduta da revisão

Com a finalização do planejamento de revisão é introduzido a conduta da revisão, no qual, nessa fase é definido uma plataforma para se explorar as publicações sobre o tema de MCLP. A plataforma escolhida foi *Web of Science*, que é uma plataforma bastante utilizada no meio acadêmico para indexação de citações científicas on-line, fundamentada em assinaturas que são

originalmente produzidas pelo *Institute for Scientific Information* (ISI) depois alimentado pela *Clarivate Analytics*. Em outras palavras, a plataforma possibilita o acesso a vários bancos de dados, facilitando assim a pesquisa e a exploração de diversas conexões entre os artigos.

Para que a pesquisa tivesse um embasamento forte e de uma qualidade relevante, foi definido que apenas seriam escolhidos os artigos completos e publicados em repositórios de grande impacto para área acadêmica. Dessa forma, a revisão da literatura iria possibilitar uma ampla visão dos principais pontos que estão sendo trabalhados pelos principais autores.

Seguindo essa diretriz, foi definido dois campos de busca na plataforma do *Web of Science* que podem ser visualizados no quadro 1. O primeiro campo foi utilizado para especificar o tema principal a ser explorado, nesse caso, o modelo *Maximum Covering Location problem* (MCLP) e suas demais variações de como o tema é chamado. O segundo campo foi usado para definir uma das principais áreas na qual é aplicado os modelos MCLP, como o problema de localização, segurança pública, instalações policiais, instalações de emergência, unidades policiais e modelos de localização. Alguns operadores foram utilizados para não perder nenhum artigo da área da pesquisa. Foi usado “OR” quando o campo necessitava buscar mais de um tema em paralelo, ou seja, “tema 1” OU “tema 2”. E para especificar que deveriam ser procurados artigos que coexistisse com mais de um tema no mesmo artigo, foi usado o operador “AND”, ou seja, “tema 1” e “tema 2”.

Quadro 1 – Palavras chaves e Operadores

Palavras-Chaves E Operadores	
Campo 1	"MAXIMAL COVERING LOCATION PROBLEM"OR"MAXIMUM COVERING LOCATION PROBLEM"OR"MCLP"
	"Campo 1"AND"Campo 2"
Campo 2	"FACILITY LOCATION PROBLEM"OR"SAFETY"OR"PUBLIC FACILITIES"OR"POLICE FACILITIES"OR"EMERGENCY FACILITIES"OR"POLICE UNITS"OR"LOCATION"OR"FACILITIES"OR"FACILITY LOCATION"OR"MODELS"

Fonte: O Autor (2019)

Durante a primeira busca na plataforma foi encontrado 210 publicações contendo os tópicos supracitados referentes ao campo 1 e o campo 2. Após aplicar o filtro para conter apenas artigos completos, a base de dados reduziu para 144 artigos que estão entre um intervalo de tempo de 1976 até 2019. Porém esse refino ainda não é o suficiente, se faz necessário, analisar se entre esses 144 artigos existe algum artigo que não corresponde a problemática da revisão. Logo, foi realizado essa análise através do resumo de todos esses artigos restantes para eliminar os artigos não correspondentes, chegando em um número final de 120 artigos.

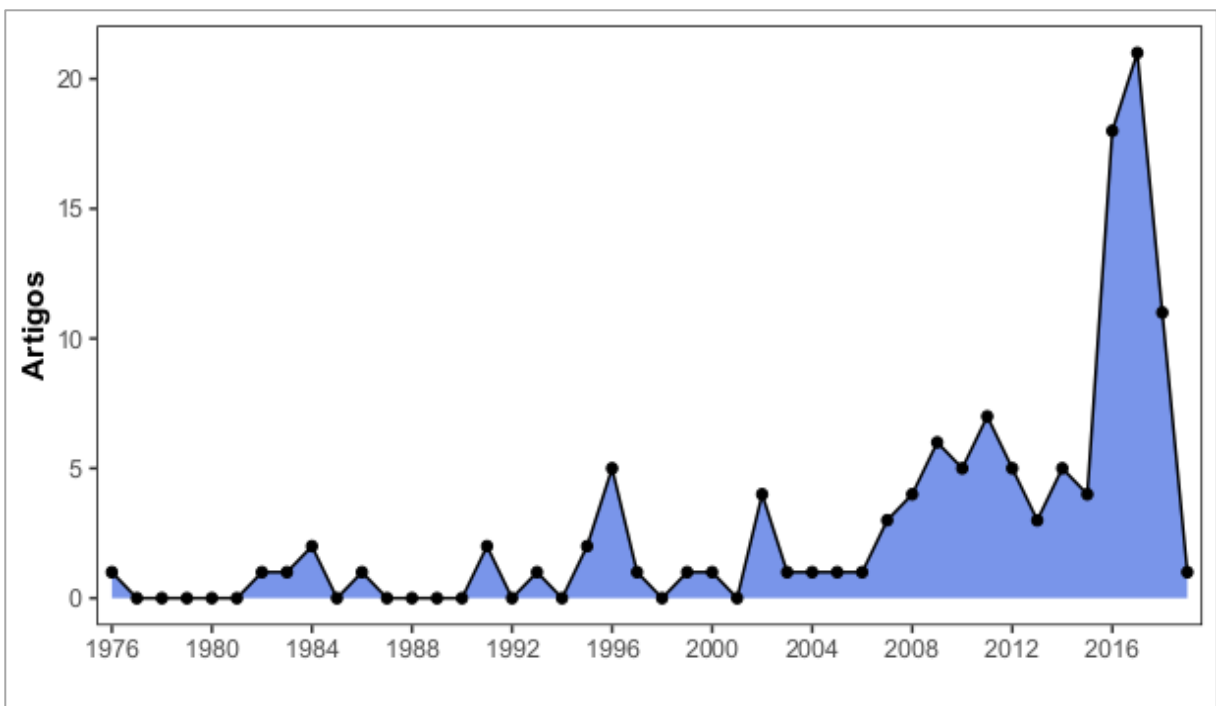
3.1.3 Documento de revisão

Chegando na terceira fase o primeiro passo é filtrar as informações necessárias dos 120 artigos para a construção o documento de revisão da literatura. As principais informações que devem ser retiradas dos artigos para responder as questões levantadas na primeira fase da metodologia, são, os números de autores, trabalhos publicados, área de atuação, citações, local de publicação, país de origem e conexões entre autores.

Porém, para que esse processamento de informação seja eficaz, é utilizado o *Software RStudio* com a utilização de uma biblioteca voltada para o processamento de revisão literária. Essa biblioteca é chamada de *Bibliometrix* e através de suas funções internas é possível realizar o processamento dos artigos e gerar gráficos e estudos de correlações entre os autores e as citações (ARIA e CUCCURULLO, 2017).

Com base no *Bibliometrix* podemos iniciar a construção do relatório e de suas discussões. Nesse sentido, é possível observar no Gráfico 3 que o modelo *Maximal Covering Location Problem* (MCLP), teve um aumento bastante significativo ao longo tempo. Saindo de 1 artigo publicado em 1976 para um pico de 21 artigos em 2017.

Gráfico 3 – Publicação anual científica sobre o modelo MCLP



Fonte: O Autor (2019)

Esse aumento substancial de publicações sobre o modelo MCLP é devido as necessidades do mercado, que cada vez mais, evoluindo quando se trata de localizar recursos escassos (REVELLE e EISELT, 2005). Concomitantemente ao longo desse período, o avanço do poder

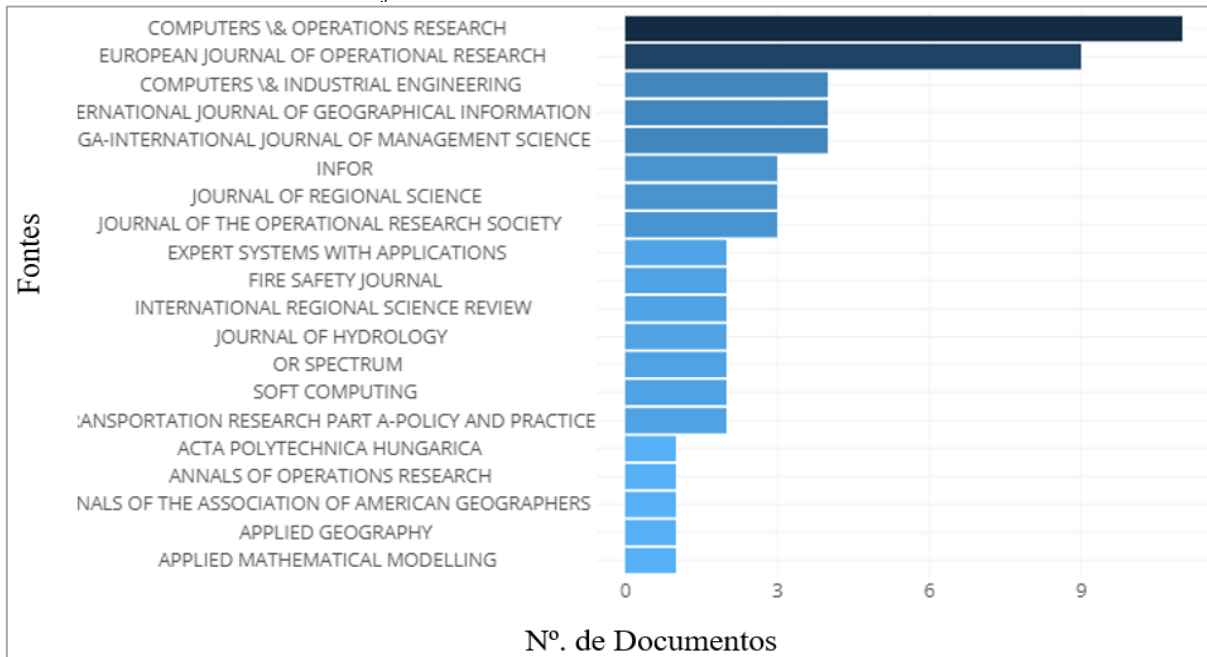
computacional foi extremamente significativo (MURRAY, 2016), facilitando a implementação do modelo em diversas plataformas computacionais.

O modelo que é voltado a solucionar o problema de localização de instalações, evoluiu muito com relação a sua aplicabilidade. Church e ReVelle (1974) afirmam que o MCLP originalmente proposto, seria particularmente mais valioso em áreas de serviços emergenciais. Mas sua aplicação se estendeu muito além, com o MCLP sendo frequente utilizado para apoiar o gerenciamento de recursos naturais, publicidade, fabricações, processamento de localizações e etc. (MURRAY, 2016).

Foi possível observar que o modelo MCLP foi aplicado em diversas áreas diferentes tanto para o setor público como para o privado. Com pequenas adaptações, o modelo foi aplicado em localização de base de helicóptero (GARNER; BERG, 2017), bases de serviços médicos e equipamentos médicos (VATSA; JAYASWAL, 2016; PULVER *et al.*, 2016), bases de ambulâncias (RØISLIEN *et al.*, 2017; YIN e MU, 2012), base de bombeiros (WANG *et al.*, 2016), base de resgate (ZHOU *et al.*, 2019), instalações postais, delegacias (MCCALL e QIU, 2007; GUSMÃO *et al.*, 2019), antenas de TV, antenas de celular (YIN e MU, 2015), antenas de RFID (OZTEKIN *et al.*, 2010), redes de banda larga (GRUBESIC e MURRAY, 2002), reservas ambientais (LORAAMM e DOWNS, 2016; SNYDER e HAIGHT, 2016; CHURCH *et al.* 1996), torres de vigilância (BAO *et al.*, 2015), sirenes de emergência (CURRENT e OKELLY, 1992), sensores (WANG *et al.*, 2017) de segurança (MURRAY *et al.*, 2007), anúncios em outdoor (HUANG *et al.*, 2019), paradas de ônibus (OTTO e BOYSEN, 2014) e etc. Tudo isso reflete, o quanto, o modelo é sólido para solucionar esse tipo de problema por mais antigo que o mesmo seja.

Já no Gráfico 4, temos a relação das 20 principais fontes dos artigos analisados durante a revisão da literatura. Na qual, a COMPUTERS & OPERATIONS RESEARCH está em destaque com 11 artigos em seu banco de dados nessa área de pesquisa, seguidos da EUROPEAN JOURNAL OF OPERATIONAL RESEARCH com 9 artigos e COMPUTERS & INDUSTRIAL ENGINEERING, GEOGRAPHICAL INFORMATION, MANAGEMENT SCIENCE com 4 artigos cada um. E juntos são responsáveis por mais de 26% das publicações em análise. Valor esse bastante significativo dado que foram analisados 80 repositórios diferentes, todos de elevada qualidade e impacto na área acadêmica.

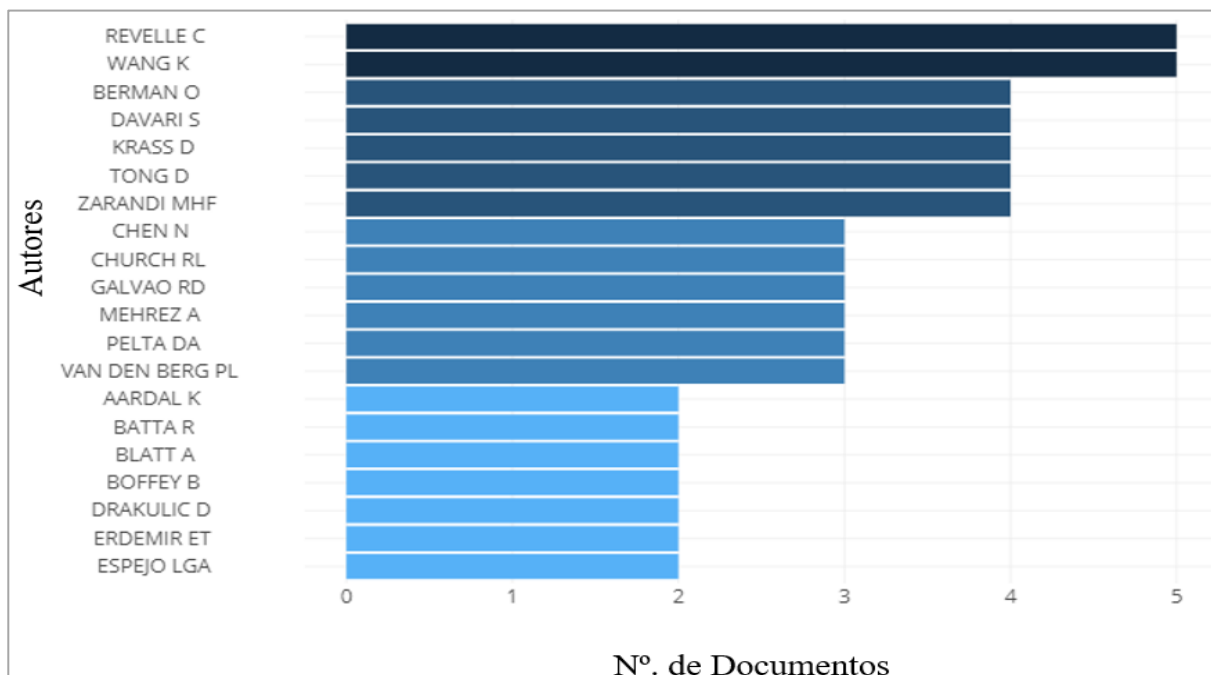
Gráfico 4 – Fontes analisadas na revisão literária



Fonte: O Autor (2019)

Seguindo essa analogia, é plausível realizar esse mesmo tipo de análise para os autores. Essa análise é importante pois é demonstrado quem são os principais autores que contribuíram para o avanço do modelo MCLP. Dessa forma, no Gráfico 5 é ilustrado a relação de publicações dos 20 principais autores, e é notória a presença de Revelle e Wang que estão na liderança com 5 artigos publicados cada um, segundo os princípios da metodologia abordada nessa revisão.

Gráfico 5 – Relação de publicações por autor

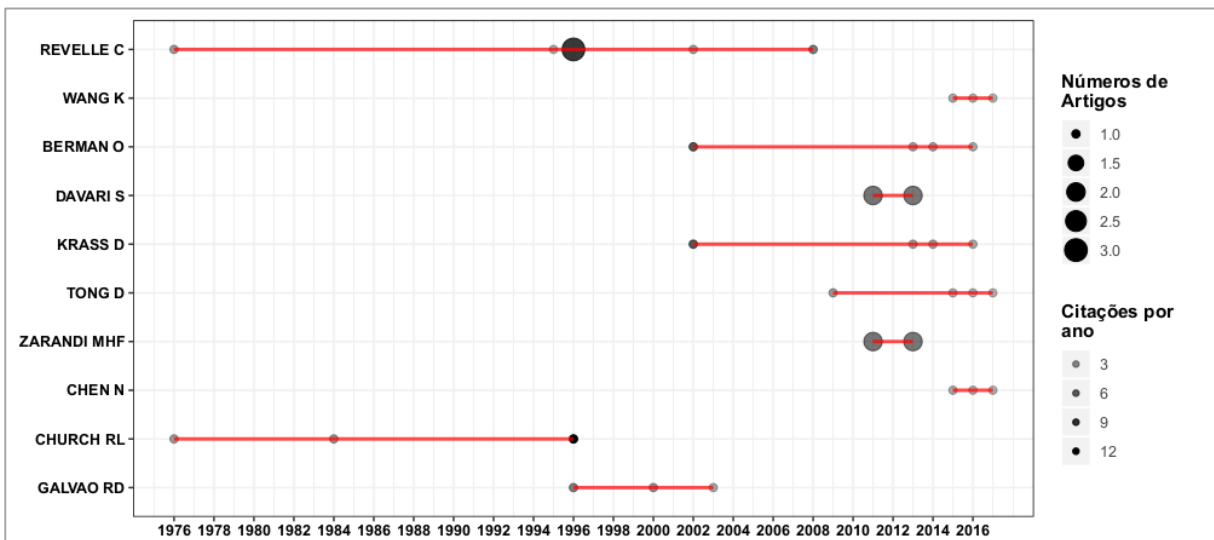


Fonte: O Autor (2019)

Nesse momento, vale destacar que o principal artigo de Church e Revelle, no qual, apresenta a criação do modelo MCLP em 1974 não está incluído nos dados do Gráfico 5. O artigo de criação do modelo não consta na base de dados da *Web of Science*. Isso significa que Revelle tem mais uma publicação, ficando com 6 artigos publicados e se tornando o autor de maior relevância na área, e Church com 4 artigos.

Logo, para melhor entender o comportamento dos principais autores na área de MCLP foi desenvolvida uma análise temporal sobre suas publicações. No Gráfico 6 temos a produção dos autores ao longo do tempo, relatando os espaços de tempo entre as publicações de cada autor. Um ponto bastante importante dessa análise é que ao avançar do tempo o intervalo de publicações diminuiu, mostrando assim, o interesse dos autores pelo tema e corroborando com o crescimento elevado de publicações ao longo do tempo apresentado na Gráfico 3.

Gráfico 6 – Produção dos autores ao longo do tempo



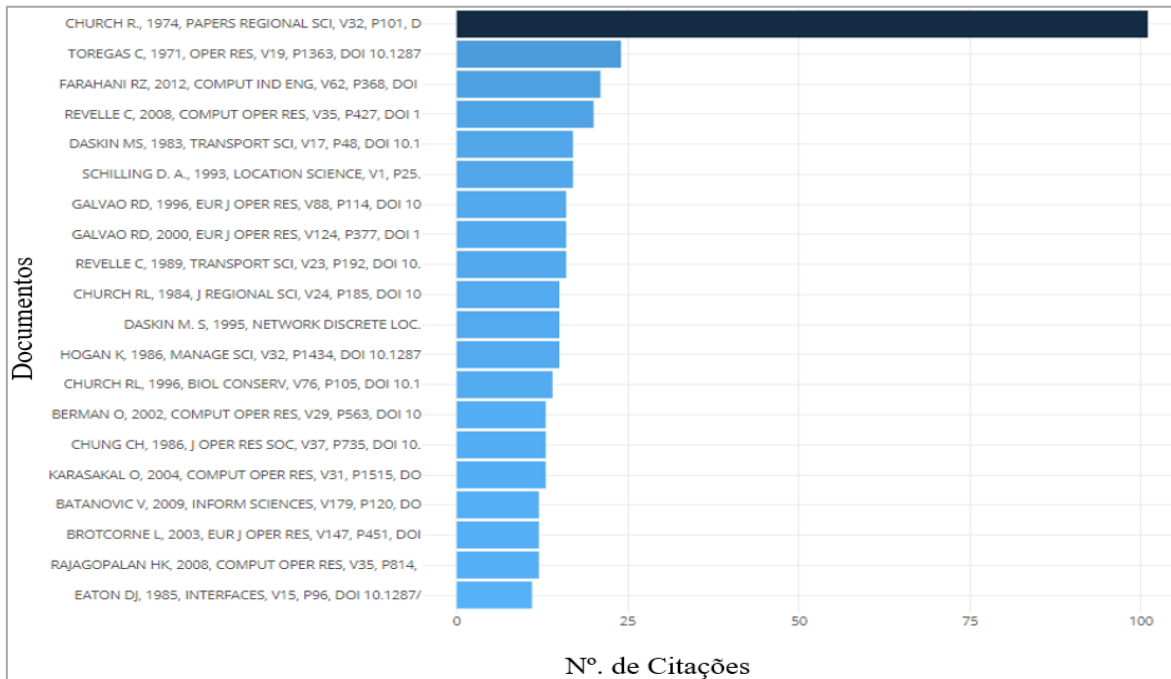
Fonte: O Autor (2019)

É notório que para um artigo tenha relevância no meio acadêmico o mesmo deve ser bem referenciado. Nesse sentido, dentre os 120 artigos analisados entre 1976 – 2019 nesta revisão, podemos encontrar que para a criação desses artigos foram necessários 301 autores, existindo uma média de 3,09 coautores por documentos, contendo apenas 5 documentos de autoria única. E foi observado uma média de 19 citações por documento, número bem considerado dado ao espaço de tempo que foi analisado.

E quando se trata do modelo MCLP, Church e Revelle como os criadores do modelo acabam se tornando a principal referência para qualquer outro trabalho que for ser executado na área. Essa informação pode ser confirmada no Gráfico 7, que apresenta as referências mais

citadas dentro do conjunto de artigos analisados e em primeiro lugar se encontra Church R. com 101 citações.

Gráfico 7 – Referências mais citadas



Fonte: O Autor (2019)

Para definir quem são os países que estão à frente das pesquisas sobre MCLP é desenvolvida uma base de dados com a origem dos trabalhos contidos nessa revisão da literatura e desenvolvido um mapa-múndi que pode ser observado na Figura 9.

Nesse sentido, temos o número de documentos produzidos em cada país e em destaque podemos citar os Estados Unidos, China e Canadá com respectivamente 66, 23 e 17 documentos publicados na área. É importante frisar como o tema tem impacto ao redor do mundo, se estendendo por mais 28 países, mostrando como é forte o interesse pelo modelo voltado a problemática de localização de instalações.

Entretanto não podemos tirar uma afirmação sobre o porquê um país tem mais estudos nessa área do que outro, com apenas essa base de dados. Porém, é plausível destacar que as condições financeiras do país e os investimentos voltados a pesquisa, são fatores bastante relevantes para um aumento expressivo de publicações em um país.

O estudo da revisão da literatura mostrou-se de grande relevância, por identificar como o modelo MCLP se comportou na literatura desde sua criação. Servindo como uma base mais sólida para o desenvolvimento desta pesquisa, principalmente ao demonstrar as oportunidades

de aplicações e desenvolvimento de estudos com modelagens matemáticas consistentes e flexíveis.

Figura 9 – Produção de documentos por país



Fonte: O Autor (2019)

Logo, foi possível identificar o expressivo crescimento do assunto quanto ao número de publicações, o grau de participação de diversos países no estudo e no desenvolvimento de novas aplicações, além da expressividade que o modelo tem em um cenário mundial.

3.2 Justificativa do uso do modelo

É notório que, o MCLP é uma técnica de localização muito poderosa que de fato complementa o julgamento dos especialistas que antes realizavam a alocação das instalações com apenas seus conhecimentos empíricos (CHURCH e REVELLE, 1974). O modelo oferece uma visão extensa do problema que contribui para uma rápida tomada de decisão, alocando de forma otimizada as instalações e realizando o melhor uso possível do espaço e dos recursos disponíveis (ERDEMIR *et al.* 2010).

Corroborando com essa afirmação, Loraamm e Downs (2016) apresentam que a maior vantagem de usar a abordagem do modelo MCLP, está na clareza e simplicidade em que os resultados são apresentados, apontando os locais ideais para as instalações e facilitando o cruzamento de diferentes cenários de cobertura para o decisor. Grubestic e Murray (2002) complementam com a vantagem da minimização dos custos que podem ser alcançadas com a

utilização do modelo, mediante a redução da quantidade de instalações que podem ser alocadas para atender a cobertura exigida pelo decisor.

Com base nesses aspectos, o modelo MCLP foi amplamente aplicado em diversos setores, principalmente os de caráter emergencial. Tendo um grande salto em suas aplicações nos últimos anos como foi analisado na seção anterior, mostrando assim, o quanto o modelo é sólido com relação a sua estrutura matemática. Segundo Murray (2016) existem mais de 1.550 citações até hoje na literatura acadêmica sobre o tema, corroborando que o modelo resistiu ao teste do tempo, podendo ser tão relevante ou mais nos dias de hoje do que quando foi publicada pela primeira vez.

Entretanto, através da revisão da literatura foi observado que na maior parte dos casos o modelo MCLP foi aplicado de maneira isolada, ou seja, sem estar atrelado a outros modelos que possivelmente maximizasse a sua capacidade. Evidentemente, em alguns dos casos o modelo foi adaptado para a problemática em que estava sendo estudada, porém não foi utilizada em conjunto com outros modelos matemáticos.

Segundo Silva *et al.* (2013), é possível fazer combinações de modelos para obter melhores desempenhos nos resultados finais, quando se trata de aplicações de modelos de localização de instalações. Nesse sentido, o autor apresenta que os locais candidatos para se alocar uma instalação podem ser encontrados por modelos matemáticos e depois ser utilizado um modelo de localização de instalações para selecionar o melhor local entre esses locais pré-definidos, citando ainda como exemplo o modelo *k-means*.

Corroborando com essa afirmação, Risso *et al.* (2015) definem que o modelo *k-means* é consistente principalmente para os modelos de localização de *covering problem*, visto que o modelo busca encontrar centroides em clusters. E é algo que não foi encontrado na revisão da literatura para o modelo MCLP. Gerando assim, um campo de estudo que pode ser explorado para maximizar ainda mais os resultados no momento da alocação de uma nova instalação.

Nesse sentido, é plausível o desenvolvimento de novos estudos nos dias atuais com o modelo MCLP, especialmente por existir novos campos em abertos que podem ser explorados. E que se adequam perfeitamente nesta pesquisa, por se tratar de um problema de localização de um serviço emergencial em um departamento público e que precisa ser maximizado.

4 METODOLOGIA

Esta pesquisa busca propor um sistema de informação gerencial (SIG), baseado em modelos e técnicas computacionais, tais como o K-means e o MCLP, para dar suporte a análise e visualização de dados sobre o arranjo de instalações policiais na cidade do Recife-PE. Nesse sentido, este capítulo será apresentado todos os procedimentos metodológicos utilizados nesta pesquisa, apresentando todas as fases, desde o método de pesquisa, a caracterização do estudo, bem como a escolha dos procedimentos técnicos utilizados para aplicação do modelo MCLP.

4.1 Classificação da pesquisa científica em engenharia de produção

Toda pesquisa pode ser classificada mediante a seus componentes internos. Nesse sentido, uma pesquisa na área de engenharia de produção pode ser classificada em sua natureza, objetivos, abordagem e o método utilizado (TURRIONI e MELLO, 2012; GIL, 2010)

Essa classificação pode ajudar aos leitores a identificar com rapidez se a pesquisa em questão está de acordo com a sua busca. Logo, segundo Turrioni e Mello (2012) esta pesquisa pode ser classificada em natureza aplicada. No qual, os resultados da pesquisa podem ser aplicados ou utilizados na prática para solucionar um problema real.

Quanto aos objetivos dessa pesquisa Turrioni e Mello (2012) classificam com caráter exploratório, buscando uma maior familiaridade com o problema, através do levantamento de informações para sugerir uma melhoria no problema. Já a classificação da abordagem dessa pesquisa é combinada, no qual, envolve tanto uma análise de dados quantitativos quanto dos dados qualitativos (GIL, 2010).

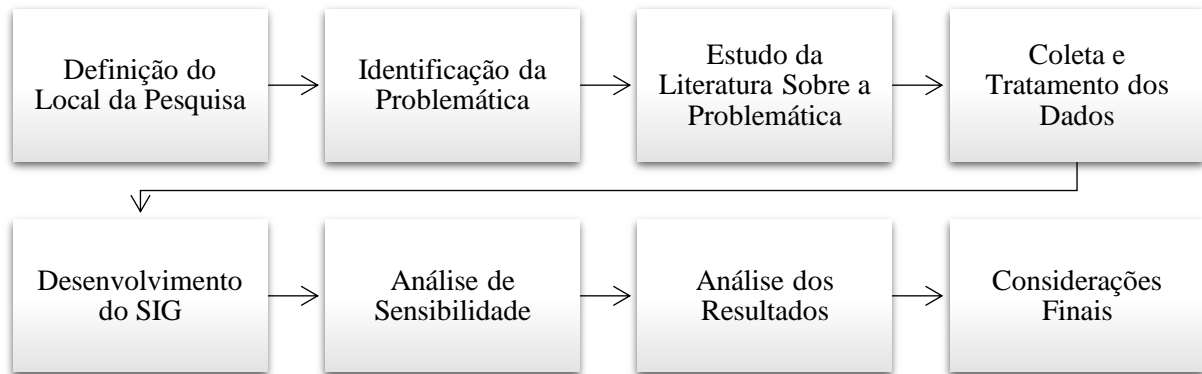
Por fim a classificação do método utilizado para essa pesquisa é o estudo de caso, envolvendo um estudo aprofundado e exaustivo de alguns objetos do problema para que se permita o seu amplo e detalhado conhecimento (TURRIONI e MELLO, 2012). Ao realizar essa classificação, é necessário apresentar como o trabalho foi iniciado e dividido.

4.2 Delineamento com as fases da pesquisa

A diretriz na qual foi estruturada todas as fases desta pesquisa é apresentada no Fluxograma 2. Segundo Gil (2010) essa diretriz em formato de fluxograma recebe um nome específico na literatura acadêmica de Delineamento da Pesquisa. Esse delineamento visa apresentar a sequência na qual foi planejada e dirigida a pesquisa, possibilitando que outros pesquisadores possam segui-lo, caso busquem replicar o projeto ou parte dele em um trabalho

futuro. Nesse sentido, as próximas subseções apresentam e descrevem a importância de cada fase.

Fluxograma 2 – Delineamento da pesquisa



Fonte: O Autor (2019)

4.2.1 Definição do local da pesquisa

O local de estudo que foi escolhido para esta pesquisa, foi a cidade do Recife que é a capital de Pernambuco. Esse local do estudo foi escolhido por se tratar de uma região que vem crescendo os seus índices de criminalidade, como foi descrito na seção 2.1.1 e 2.1.2.

Esse aumento dos índices criminas, pode ser analisado no 12º ANUÁRIO DE SEGURANÇA PÚBLICA que publicado em 2018. No qual, consta que apenas para os CVLI de Recife entre os anos de 2016 a 2017, houve um aumento de 22% de casos.

4.2.2 Identificação da problemática

Depois de analisar o cenário no qual a cidade do Recife-PE encontra-se, com relação a seus índices criminais, identificou-se a carência de um diagnóstico matemático e formal para o gerenciamento da cobertura proposta por cada instalação. Segundo Claro e Nickel (2002), esse modelo de diagnóstico visa evidenciar as necessidades passadas, presentes ou futuras por intermédio de pesquisas internas, a fim de identificar e descrever o problema para sugerir uma intervenção ou melhoria na problemática estudada.

Nesse sentido, a realização de um diagnóstico possibilita identificar alguns dos possíveis problemas que podem estar acontecendo com relação a cobertura prestada por cada instalação policial. Servindo como *input* para o processamento de novas medidas corretivas voltadas a localização e cobertura das instalações com relação a demanda da região.

4.2.3 Estudo da literatura sobre a problemática

Para melhor entender a problemática da pesquisa, se faz necessário a realização de um estudo da literatura, visando buscar um profundo conhecimento em problemas de localização de instalações, segurança pública e os modelos que são mais utilizados para solucionar essa problemática. Nesse sentido, a principal vantagem de se levantar um referencial teórico nessas áreas, é permitir ao pesquisador obter uma visão geral e estruturada do problema, além de se tornar indispensável para a realização das discussões e apresentações dos pontos mais relevantes da pesquisa (MARCONI e LAKATOS, 2010).

Corroborando com essa afirmação Aquino *et al.* (2008) demonstram que, esse estudo da literatura fornece a sustentação para a seleção das fontes de dados e para o estabelecimento das relações causais que irão guiar a pesquisa. Nesse sentido, para a construção do referencial teórico desta pesquisa foram utilizados como principais fontes para a coleta de artigos, jornais e revistas, os bancos de dados do portal de periódicos CAPES e os da *Web of Science*, acessados através do domínio da UFPE, que possibilita maior acesso a base de dados.

4.2.4 Coleta e tratamento dos dados

Uma das fases maior relevância do delineamento da pesquisa é a coleta e tratamento dos dados, segundo Oliveira e Moraes (1994) as pesquisas precisam ser realizadas visando a analisar a realidade e descobrir as necessidades existentes do problema. Dessa forma, se faz necessário à realização de uma coleta de dados para elencar essas informações, pois, são os ingredientes mínimos para que sejam trabalhados em busca de uma solução ou melhoria para o problema em estudo (TURRIONI e MELLO, 2012).

Uma vez feita a captação dos dados do problema, nesse caso, a localização das instalações policiais e os pontos geográficos onde ocorreram os crimes, é preciso saber o que se fazer com eles. As informações brutas raramente são uteis, e é necessário realizar um tratamento desses dados para que sejam possíveis de se manipular em um ambiente virtual (TURBAN *et al.*, 2005).

Nesse sentido, as coordenadas das instalações policiais existentes foram retiradas a partir de seu endereço e as coordenadas de onde ocorreram os crimes, foram retirados da plataforma Onde Fui Roubado. Essa plataforma foi selecionada por ser única plataforma que trabalha com informações de ocorrências de crimes e que é disponível para o público. E dessa forma, todas as informações da plataforma foram inicialmente trabalhadas em formato .xlsx (*Microsoft*

Excel) e tratadas para o formato *.txt* (*Notepad*) com separação por tabulação, no qual, foi utilizada como *input* para a leitura e processamento virtual através do *Software RStudio*.

4.2.5 Desenvolvimento do SIG

A fase de desenvolvimento do SIG, que tem como objetivo principal a implementação do modelo *Maximal Covering Location Problem* (MCLP) já descrito na subseção 2.3.3.2 e no capítulo 3, foi totalmente realizada dentro do ambiente virtual do *Software RStudio*. No qual, foi adotado o seguinte fluxograma apresentado no Fluxograma 3 para a processo de implementação e análise dos dados.

Fluxograma 3 – Implementação do modelo MCLP no Software RStudio



Fonte: O Autor (2019)

O *Software RStudio* ao iniciar, carrega automaticamente um pacote de funções básicas para que o usuário possa realizar simples manipulações de dados. Esse procedimento é realizado dessa forma para não sobrecarregar o computador do usuário, dado que no diretório do *Software* pode chegar a conter centenas de pacotes extras. Nesse sentido, quando é preciso realizar uma análise mais específica dentro do ambiente virtual do *Software RStudio*, se faz necessária a ativação das bibliotecas extras. Nesta pesquisa foi utilizada diversas bibliotecas, como: *tidyverse*, *cluster*, *factoextra*, *gridExtra*, *maxcovr*, *dplyr*, *readr*, *leaflet*, *purrr*, *ggplot2*, *shiny* e *cowplot*.

Da mesma forma, deve ser realizado a importação dos bancos de dados, contendo tanto as informações das instalações policiais como as informações dos crimes cometidos. Durante esse procedimento é realizado a leitura dados, para selecionar e alocar apenas as informações necessárias para o estudo da problemática.

Ao realizar a compilação dos dados no *Software RStudio* é possível realizar uma avaliação do cenário atual. Nesse momento, é plausível verificar a demanda total e a quantidade de instalações para suprir essa demanda. Analisando que cada instalação policial tem um raio de atuação, ou seja, uma área de cobertura, é observado a porcentagem da demanda coberta e descoberta por cada instalação, resultando assim em um diagnóstico do cenário atual vivenciado na cidade do Recife-PE.

Mediante as informações da avaliação do cenário atual, é possível começar o tratamento dos dados para encontrar os possíveis locais para alocação de instalações em pontos otimizados. Para a definição dos possíveis locais, foi utilizado uma técnica de clusterização chamada *K-means* já descrita da subseção 2.4. Onde a principal ideia é definir K centroides, para cada K *cluster* de um conjunto de dados que denotam de um padrão ou classe. Nesse sentido, no caso desta pesquisa esses padrões foram definidos por algumas características do crime cometido, como a frequência do crime e as distâncias entre os crimes, na busca por localizar a instalação em um ponto que cobrisse mais pontos levando em consideração o comportamento do crime cometido.

Obtendo o conjunto de locais candidatos, se torna factível a aplicação do modelo *Maximal Covering Location Problem* (MCLP). No caso desta pesquisa foi utilizado uma biblioteca específica no *Software RStudio* para o problema de localização descrito por Church e Reville em 1974. Essa biblioteca é chamada de *maxcovr*, que foi desenvolvida por Nicholas Tierney da Universidade Monash, localizada na Austrália. A biblioteca atualmente vem recebendo incrementos de novas funções para aumentar seu poder de processamento e facilitar o seu uso para o usuário.

Através da utilização da biblioteca *maxcovr* é encontrado o conjunto de locais otimizados e os dados voltados a quantidade de crimes cobertos por cada instalação, incluindo a porcentagem de demanda coberta e não coberta por todas instalações juntas. Nesse sentido, com os resultados otimizados foi possível desenvolver uma representação geográfica. Essa representação geográfica consiste de um mapa interativo e dinâmico com todas as coordenadas plotadas que são correspondentes aos crimes e as instalações, concomitantemente com suas áreas de cobertura. Assim, torna-se notória a compreensão dos resultados obtidos durante a aplicação do modelo.

4.2.6 Análise de sensibilidade

Segundo Andrade (2009) para extrair o máximo de informação de um modelo, o mesmo precisa de ser explorado em diversos cenários, ou seja, realizando mudanças em seus parâmetros para que seja desenvolvido uma análise de sensibilidade. Essa análise de sensibilidade foi verificada nesta pesquisa, fornecendo diferentes valores para a quantidade de instalações a serem alocadas e a quantidade de quilômetros fornecida em seu raio de cobertura.

Nesse sentido, variando a quantidade de instalações e fixando a área de cobertura é observado o ganho de atendimento da demanda com o acréscimo gradual de cada uma das

instalações. O mesmo pode ser observado fixando a quantidade de instalações e variando a área de cobertura do conjunto de instalações alocadas. Essa informação pode agregar valor ao processo de análise da decisão sobre as instalações para o tomador de decisão que verificar se a alocação de mais uma instalação é factível ou não em termos de investimento e porcentagem adicional de cobertura.

4.2.7 Análise dos resultados

A última fase do delineamento desta pesquisa é a análise dos resultados, que visa explorar todos os dados desenvolvidos durante esta pesquisa para que seja apresentado um direcionamento para a problemática. Segundo Turrioni e Mello (2012) ao analisar e interpretar os resultados, o pesquisador se move do domínio empírico para o teórico e conseqüentemente formalizando o conceito estudado.

Nesse sentido, por meio da análise dos resultados é possível gerar um conjunto de informações cruciais para uma possível tomada de decisão com mais segurança em relação aos investimentos e o adicional de cobertura efetuado por cada instalação alocada.

Em outras palavras os resultados do SIG proposto busca informa para o decisor quais serão os locais candidatos, quais serão os locais selecionados e todos os indicadores relacionados a essas escolhas. Sempre visando apresentar um arranjo de localizações que pode ser utilizado como referência para instalação de unidades policiaes, para melhorar a eficiência operacional do atendimento às ocorrências policiaes e do patrulhamento nas regiões cobertas.

5 SISTEMA DE INFORMAÇÃO GERENCIAL PROPOSTO

Este capítulo apresenta o Sistema de Informação Gerencial (SIG) desenvolvido e proposto para solucionar problemas de *Maximal Covering Location Problem* (MCLP), com o auxílio do Modelo *Kmeans* para encontrar locais candidatos. Mais especificamente, o SIG irá se comportar como uma ferramenta de análise de dados com suporte da inteligência computacional, desenvolvida para apoiar as decisões sobre a localização de instalações policiais públicas. O capítulo será dividido na apresentação da arquitetura do SIG e no design do SIG, em outras palavras, como funciona o software desde a sua entrada, processamento e saída de dados/informações.

5.1 Arquitetura do SIG

O SIG desenvolvido e proposto para solucionar o problema dessa pesquisa, tem como objetivo principal oferecer cálculos rápidos e seguros para que os tomadores de decisão possam aumentar sua produtividade no momento da alocação das instalações policiais. Tendo como diretriz, a máxima cobertura dessas instalações com relação aos crimes que estão acontecendo na região em que as mesmas serão alocadas.

Figura 10 – Arquitetura do SIG proposto.



Fonte: O Autor (2019)

Nesse sentido, os dados serão armazenados em vários bancos de dados para que seja possível analisar várias alternativas ao mesmo tempo. Logo, esses dados podem ser transformados em informações e conhecimentos através do sistema de informação gerencial.

Portanto, o decisor não vai necessitar de deter de altos conhecimentos em computação e métodos analíticos para solucionar um problema de *Maximal Covering Location Problem* (MCLP).

Na Figura 10 é apresentado a arquitetura do SIG proposto. E nesse sentido, os principais elementos do SIG são os seguintes:

- **Software:** é responsável por coordenar diversos grupos de componentes lógicos que controlam todo o computador e o sistema de processamento de dados, garantindo o seu funcionamento. Sendo os principais *softwares*, o Sistema Operacional, R e o *RStudio*, no qual, auxiliam as diversas interações do decisor para com SIG.
- **Hardware:** engloba todos os dispositivos físicos e equipamentos utilizados para o processamento da informação, ou seja, à unidade central de processamento, à memória e os dispositivos de entrada e saída de uma dada máquina. Sendo os principais *hardwares*: o computador, a internet e os diversos periféricos como a impressora, o monitor, mouse e o teclado.
- **Dados:** é responsável por agrupar e analisar os dados voltados à segurança pública de diversas áreas geográficas. Sendo os principais dados, as ocorrências da cidade do Recife que foram obtidas a partir da plataforma “Onde Fui Roubado” durante o intervalo de 2012 até 2017. Outros dados de equivalente importância, são os parâmetros exigidos pelo modelo de localização, como a cobertura máxima de atendimento e quantidade de instalações a serem alocadas, que são definidos pelo decisor.
- **Método:** É responsável por agrupar e processar conjuntamente todos os métodos para a execução do trabalho, garantindo rapidez e confiança no processamento dos cálculos no momento de encontrar as melhores soluções para o problema de localização. Nesse sentido, os principais métodos são: um modelo de processamento de problemas MCLP, um modelo de clusterização *Kmeans* e um modelo de georreferenciamento.
- **Pessoas:** são todas as pessoas que ajudam na captação e transformação dos dados em informações, garantindo que os objetivos levantados pela organização estejam sempre alinhados com sua estratégia. Nesse sentido, as principais pessoas para o SIG proposto são os decisores e os especialistas. Onde os decisores estão encarregados de manipular o SIG em busca de sugestões para alocar as instalações policiais e os especialistas que são responsáveis por trabalhar no processamento dos dados e nos modelos matemáticos envolvidos no SIG.

- **Relatório com sugestões de localização para as instalações policiais:** é responsável por informar os relatórios que são gerados pelo SIG e permitir que o decisor efetue o seu *download*. Onde os principais relatórios são voltados aos locais candidatos, locais selecionados, resultado final da alocação e análise de sensibilidade.

Dessa forma, ao executar todos esses componentes principais da estrutura do SIG, é obtido uma sinergia entre os componentes que possibilita ao decisor, desenvolver suas pesquisas exploratórias para localizar de forma otimizada as instalações policiais. Para melhor entender essa sinergia entre os componentes da arquitetura, a próxima seção tem por objetivo apresentar o design do SIG, assim como seu funcionamento.

5.2 Design e funcionamento do SIG proposto

Para entender o funcionamento do SIG na prática, é necessário navegar em sua interface seguindo todo o fluxo necessário para alocar as instalações policiais. Portanto, o fluxo tem início com o *upload* do arquivo (base de dados de crimes) e segue pelas fases de análise estatística, definição dos locais candidatos (*Kmeans*), alocação das instalações policiais (MCLP) e por fim o *download* dos relatórios, que por sua vez, contém informações sobre qualquer uma das etapas do procedimento de alocação das instalações. Nesse sentido, esta seção apresenta o design do SIG (interface) contendo todas as formas de interação que o usuário pode conseguir através desse fluxo.

Assim, o SIG proposto é dividido em 8 telas intuitivas de interação com o usuário, onde cada tela, aqui será chamada de página. Logo, temos a página de Arquivos, Dados, *Kmeans*, Locais Candidatos, Alocação (MCLP), *Download*, Sobre e Configurações. E é bastante importante frisar, que a interface do SIG ainda está em desenvolvimento e algumas palavras que no português precisam de um caractere especial como “ç” e os diversos acentos, ainda não foram implementados completamente ao algoritmo de interface. Dessa forma, algumas poucas palavras dentro da interface serão apresentadas sem esses caracteres especiais.

Nesse sentido, a Figura 11 apresenta a primeira página do SIG, Arquivo. É possível observar que a tela é dividida em três ambientes, uma barra cinza, uma barra branca e um ambiente de trabalho. Onde a barra cinza, é o menu do SIG e nela é possível encontrar um campo de busca e todas as páginas que podem ser acessadas pelo usuário, grifando em destaque a página na qual está selecionada para apresentação, que nesse caso é a página de Arquivo.

A barra branca da Figura 11, é utilizada para importar a base de dados do usuário, no qual, é dividida em quatro segmentos, “Escolha o Arquivo”, “Separador”, “Citar” e

“Exibição”. No segmento “Escolha o Arquivo”, o usuário pode utilizar do botão “Procurar” para navegar em seu computador e escolher o seu banco de dados que vai ser importado para o SIG. Ao importar o banco de dados, será apresentado ao lado e abaixo do botão “Procurar” o nome do arquivo importado e os *status* do *upload* respectivamente. E caso o banco de dados já tenha as colunas nomeadas, é assinalado a caixa de seleção “Cabeçalho”, caso contrário, é deixado desmarcado e o próprio programa irá nomear as colunas.

Figura 11 – Página Arquivo do SIG proposto

The screenshot shows the 'Escolha o Arquivo' (Choose File) section of the software. It includes a search bar with the text 'banco_de_dados_ap.txt' and an 'Upload complete' button. Below this, there are settings for 'Cabeçalho' (checked), 'Separador' (Comma, Tab, etc.), and 'Citar' (None, Double quote, Single quote). The 'Exibição' (Display) section has 'Poucos' (Few) selected. The main table, titled 'Arquivo Importado', displays the following data:

id	endereco	lat	long	valor_prejuizo	descricao	tipo_assalto_id	sexo	munic
83046	PE-008, 6397 - Imbiribeira, Recife - PE, Brasil	-8.10	-34.91	20.00		2	1	
83055	Pte. Estácio Coimbra, 106-304 - Madalena, Recife - PE, 50720-535, Brasil	-8.06	-34.90	200.00	Dois homens, um armado e outro com a mochila recolhendo todos os celulares dos passageiros.	2	1	
82982	Condomínio do Edifício Antibes - R. Prof. Antônio Coelho, 934 - Várzea	-8.05	-34.95	1500.00	Indo pra UFPE por volta das 15:00h fui abordado por um homem em	2	1	

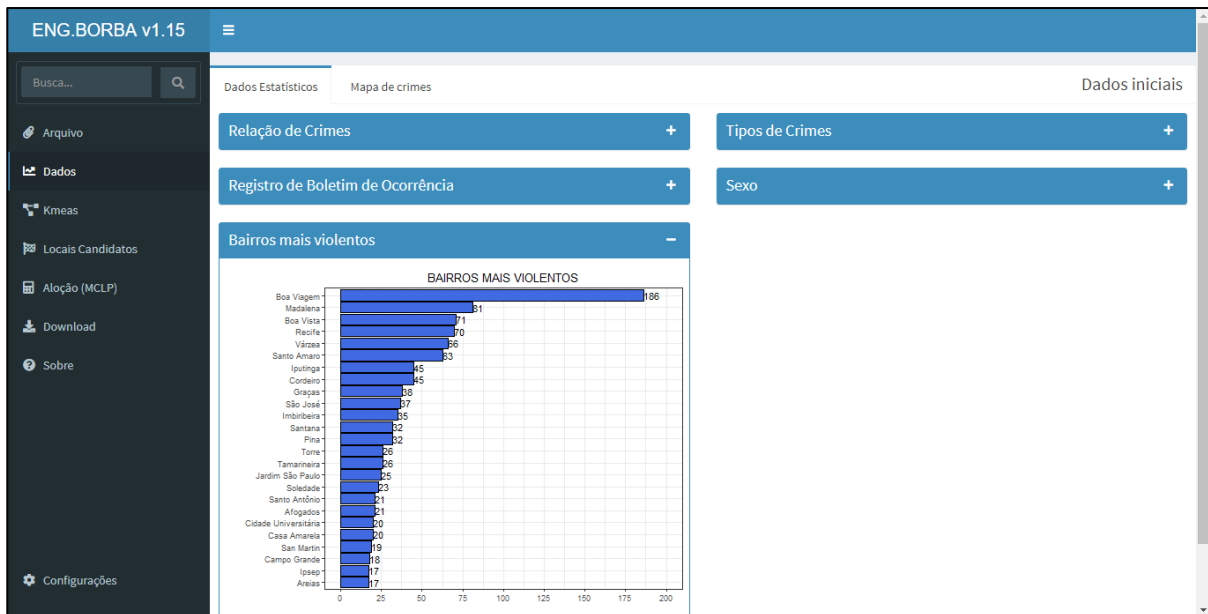
Fonte: O Autor (2019)

Diversas são as formas em que os bancos de dados são produzidos e organizados. Em alguns casos, os dados são organizados por vírgula e em outros por tabulação, alguns podem conter textos dentro de aspas duplas ou mesmo caracteres dentro de aspas únicas. Visando essa problemática, foi desenvolvido os dois segmentos, “Separador” e “Citar”, que contém caixas de seleção para que o usuário possa adequar da melhor forma seu banco de dados dentro do SIG.

Já o último segmento da barra branca na Figura 11 chamado de “Exibição”, trabalha em conjunto com o ambiente de trabalho ao lado, no segmento “Arquivo Importado”. Fornecendo para o usuário, uma prévia de como seu arquivo está sendo organizado para importação dentro do SIG. Ou seja, o usuário pode selecionar a opção “Poucos”, para analisar apenas as primeiras linhas do banco de dados, ou mesmo, a opção “Tudo”, para ter acesso a todo o banco de dados. E caso essa base de dados seja muito extensa, pode ser utilizado a barra de rolagem do SIG para navegação e análise de todos os dados importados.

Na Figura 12, é observado que na barra cinza, a página Dados está selecionada e seu conteúdo é apresentado no ambiente de trabalho ao lado. Esse ambiente de trabalho da página Dados é dividido em duas guias, que podem ser observadas no canto superior esquerdo do ambiente de trabalho, chamadas de “Dados Estatísticos” e “Mapa Crimes”. Para identificar qual guia está aberta dentro da página selecionada, é apresentado uma linha azul acima do nome da guia, no caso da Figura 12 é destacado que a guia aberta é a “Dados Estatísticos”.

Figura 12 – Página Dados do SIG proposto, guia Dados Estatísticos



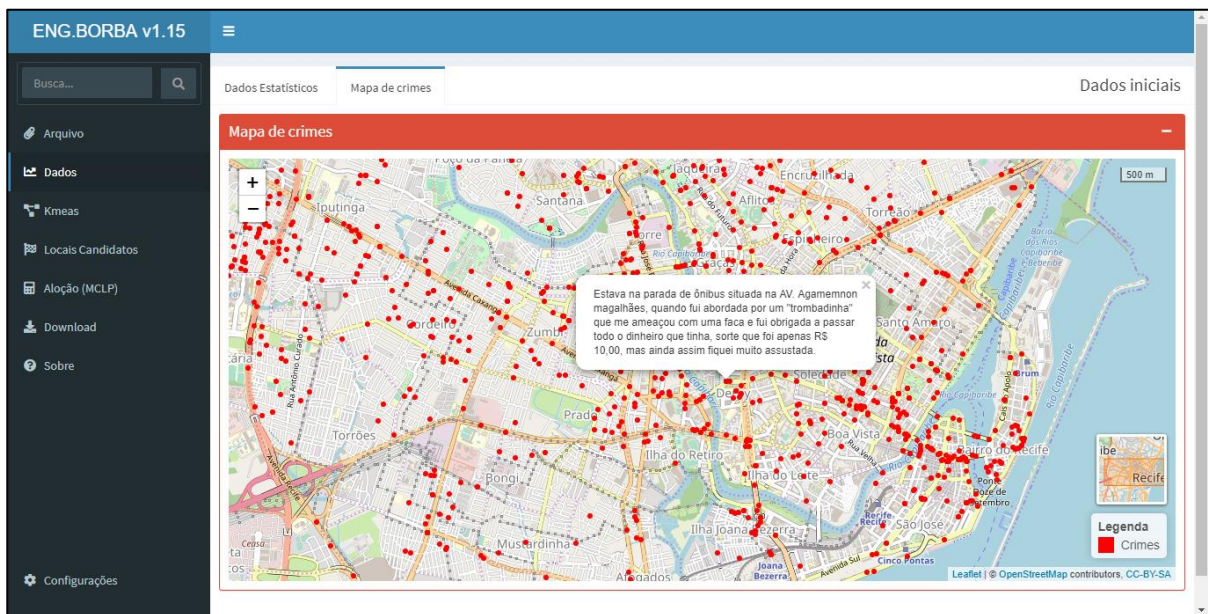
Fonte: O Autor (2019)

Nesse sentido, a guia de “Dados Estatísticos” é responsável por realizar o processamento dos dados iniciais que foram importados para o SIG. Ou seja, é possível realizar uma análise estatística dos dados, caso o banco de dados contenha informações sobre o ano em que o crime foi cometido, os diferentes tipos de crimes, se a vítima registrou ou não o boletim de ocorrência (BO), sexo da vítima e o bairro onde ocorreu o crime.

Para observar cada uma dessas análises supracitadas, a guia “Dados Estatísticos” fornece cinco segmentos de caixas, “Relação de Crimes”, “Tipos de Crimes”, “Registrou Boletim de Ocorrências”, “Sexo” e “Bairros mais Violentos”. Cada caixa contém um sinal de adição (+) em seu canto direito e ao ser clicado no sinal, o gráfico contendo a análise estatística da caixa correspondente é aberto. Nesse sentido, é observado na Figura 12 que a caixa “Bairros mais Violentos” está aberta e apresenta um gráfico contendo as informações pertinentes ao banco de dados, facilitando para o usuário, a sua interpretação sobre o assunto. A qualquer momento essa caixa pode ser minimizada clicando no sinal de subtração (-) que surge no mesmo lugar do sinal de adição após o clique.

Na Figura 13 é apresentado a página Dados, mais especificamente o ambiente de trabalho com a guia “Mapa Crimes”. Essa guia é responsável por apresentar uma georreferência dos crimes que constam no banco de dados importado para o SIG. No ambiente de trabalho da guia “Mapa Crimes” é encontrado um segmento de caixa que contém um mapa interativo e extremamente detalhado sobre todas as ocorrências.

Figura 13 – Página Dados do SIG proposto, guia Mapa Crimes



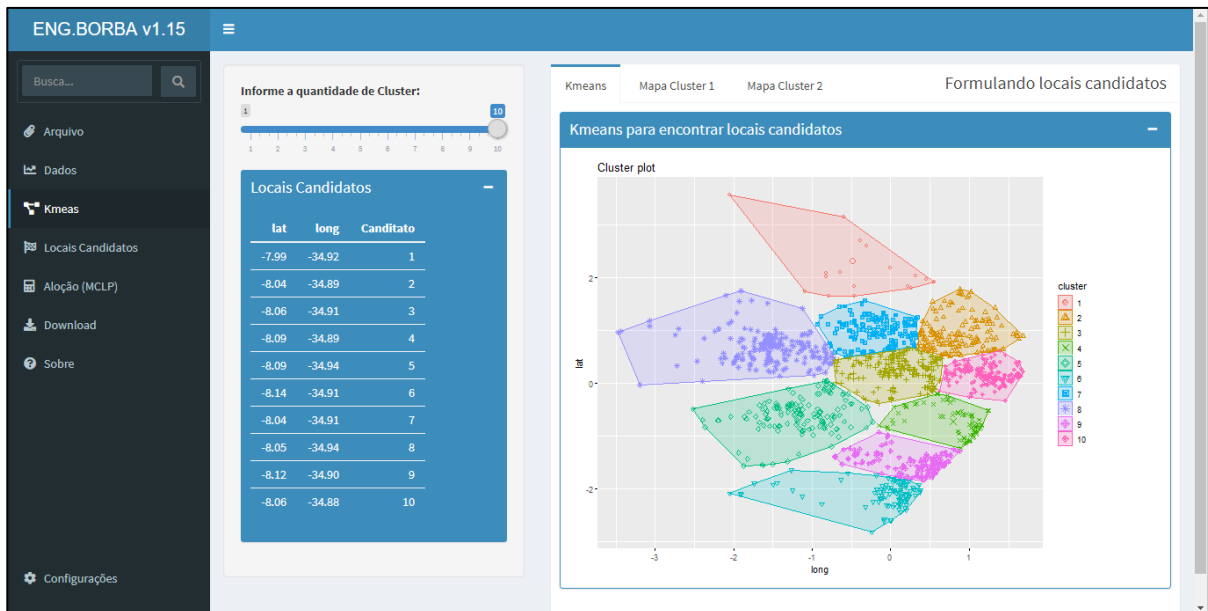
Fonte: O Autor (2019)

Além de plotar no mapa as coordenadas onde ocorreram os crimes, é possível observar no canto superior direito do mapa, a escala de distância no qual o mapa está sendo representado, mediante a altitude em que são observados os crimes. Nesse sentido, no canto superior esquerdo do mapa é possível observar dois botões com sinais de adição e subtração, que possibilita o usuário mudar a altitude em que é observado os crimes para mais e para menos respectivamente. Esses botões podem ser substituídos pela barra de rolagem do *mouse* que também pode ser utilizado para movimentar o mapa para qualquer uma das direções.

No canto inferior direito do mapa é apresentado a legenda do mapa e um mini mapa que auxilia o usuário na navegação do mapa maior. Esse auxílio é mais utilizado quando se está observando os crimes em uma altitude muito baixa e é preciso de uma visão geral da região. Outro recurso de bastante ajuda no mapa, está no ponto em que é plotado o crime. Ao clicar no ponto que representa o crime, é aberto um *pop-up* que consta uma rápida descrição do crime ocorrido relatado pela própria vítima. Um exemplo desse *pop-up* é claramente observado na Figura 13.

Com base nas informações supracitadas do design do SIG, na Figura 14 torna-se notório que é apresentado a página Kmeans e que seu ambiente de trabalho é dividido em três guias, “Kmeans”, “Mapa Cluster 1”, “Mapa Cluster 2” e a guia que está aberta é a “Kmeans”. A barra branca da Figura 14 é dividida em dois segmentos, um *input* chamado de “Informe a quantidade de Cluster:” e o outro, um *output* chamando de “Locais Candidatos”.

Figura 14 – Página Kmeans do SIG proposto, guia Kmeans



Fonte: O Autor (2019)

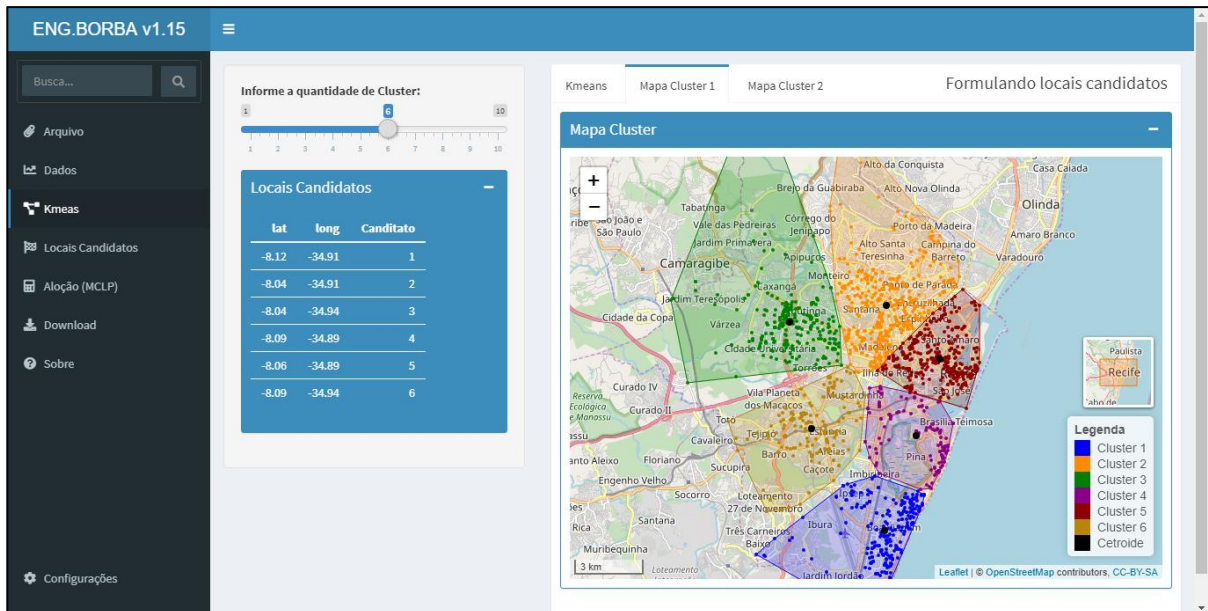
No segmento “Informe a quantidade de Cluster:”, o usuário pode manipular uma barra corrediça graduada. E dessa forma, é informando ao SIG a quantidade de cluster que o usuário deseja particionar o seu banco de dados de crimes. Os crimes são agrupados em clusters através do modelo *kmeans*, mediante a semelhança da região onde os crimes estão ocorrendo com maior frequência. Nesse sentido, com a ferramenta *kmeans* é possível encontrar o centroide de cada um dos clusters e trabalhá-los como um local candidato para se alocar uma instalação policial.

Esses centroides encontrados, são apresentados na barra branca da Figura 14 em forma de tabela no segmento de “Locais Candidatos”. Nesse segmento, o centroide de cada cluster é representado por uma coordenada geográfica, que pode mudar automaticamente mediante a quantidade de clusters que o usuário informa. Os clusters selecionados e seus respectivos limites podem ser visualizados automaticamente através da guia “Kmeans”, que contém a plotagem de todos os pontos relacionados aos crimes através dos eixos de latitude e longitude.

Para facilitar a visualização do usuário com relação a cada um dos clusters, é adicionado as informações dos limites de cada um dos clusters, assim como, o seu centroide e os crimes que o compõem. E em seguida, os mesmos são georreferenciados em mapas cartográficos

apresentados nas guias “Mapa Cluster 1” e “Mapa Cluster 2”. Na Figura 15 é observado a guia “Mapa Cluster 1” na página Kmeans. Nessa guia, o usuário pode ter uma visão mais abrangente do procedimento de clusterização, permitindo o conhecimento da região geográfica e os crimes que fazem parte de cada cluster.

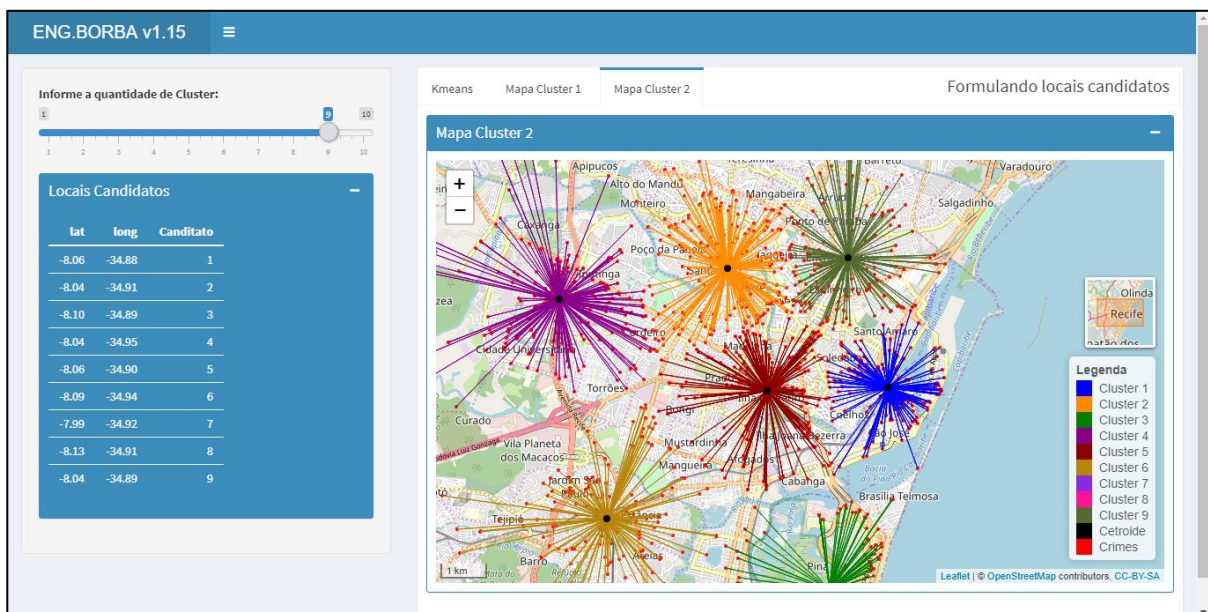
Figura 15 – Página Kmeans do SIG proposto, guia Mapa Cluster 1



Fonte: O Autor (2019)

Nesse sentido, é destacando no mapa diversos pontos pretos que representam o centroide de cada um dos clusters. Esse design de apresentação dos clusters na guia “Mapa Cluster 1”, permite que o usuário possa navegar por todo mapa e para todas as direções, principalmente quando é exigido uma visualização mais próxima ou mais afastada do cluster.

Figura 16 – Página Kmeans do SIG proposto, guia Mapa Cluster 2



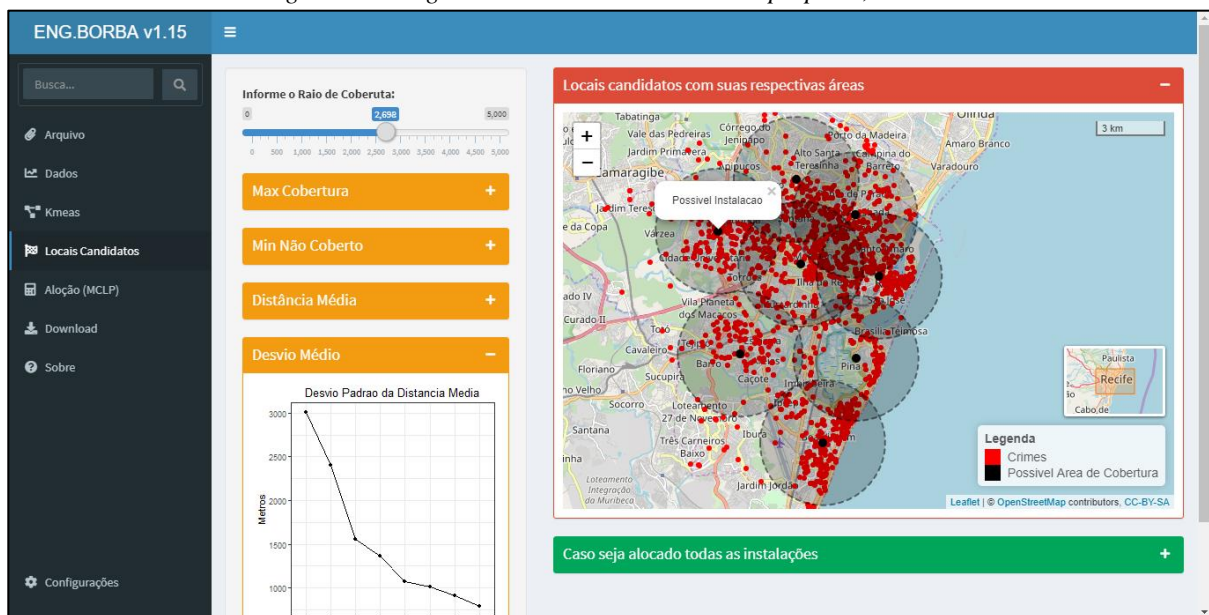
Fonte: O Autor (2019)

A Figura 16 apresenta mais uma das funções de navegação do SIG proposto. Nesse caso, é observado que a barra cinza foi ocultada, permitindo uma maior visualização do ambiente de trabalho da página Kmeans, ou seja, a guia “Mapa Cluster 2”. A barra cinza pode ser ativada e desativada em qualquer página do SIG através do botão situado na linha azul superior, ao lado direito do nome “ENG.BORBA V1.15”.

Nesse sentido, não muito diferente da guia “Mapa Cluster 1”, a guia “Mapa Cluster 2” apresentada na Figura 16 tem por objetivo representar o mapa de clusters através de uma perspectiva de distância dos crimes para o seu centroide. Em outras palavras, esse mapa permite ao usuário ter o conhecimento de quais crimes estão mais próximos e mais distantes dos locais candidatos para se instalar uma unidade policial.

Logo, com os locais candidatos definidos e o conhecimento dos crimes que estão mais próximos e mais distantes. É possível avançar para a página Locais Candidatos do SIG, que oferece uma análise dos crimes que serão cobertos por uma área de atuação proveniente de uma possível instalação que a ser alocada.

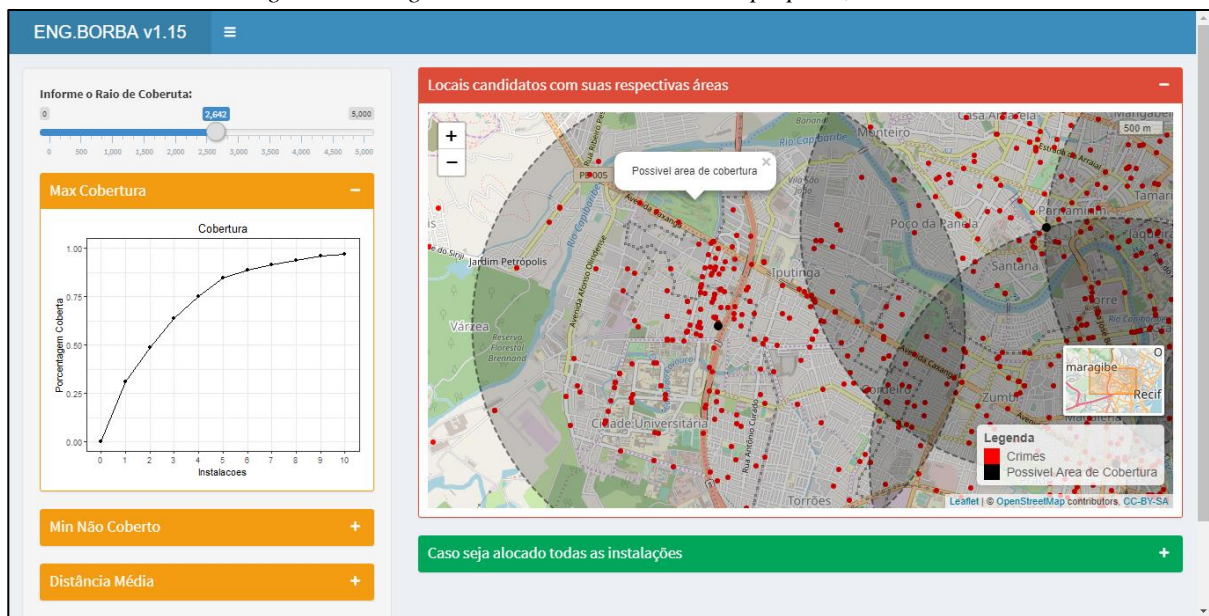
Figura 17 – Página Locais Candidatos do SIG proposto, desvio



Fonte: O Autor (2019)

Nesse sentido, na Figura 17 é apresentado a página Locais Candidatos do SIG proposto, que por sua vez, possui a divisão da barra branca em cinco segmentos. No qual, um dos segmentos é uma barra corrediça graduada, chamada de “Informe o Raio de Cobertura” e os quatro segmentos restantes, são caixas chamadas de “Max Cobertura”, “Min Não Coberto”, “Distância Média” e “Desvio Médio”.

Figura 18 – Página Locais Candidatos do SIG proposto, cobertura



Fonte: O Autor (2019)

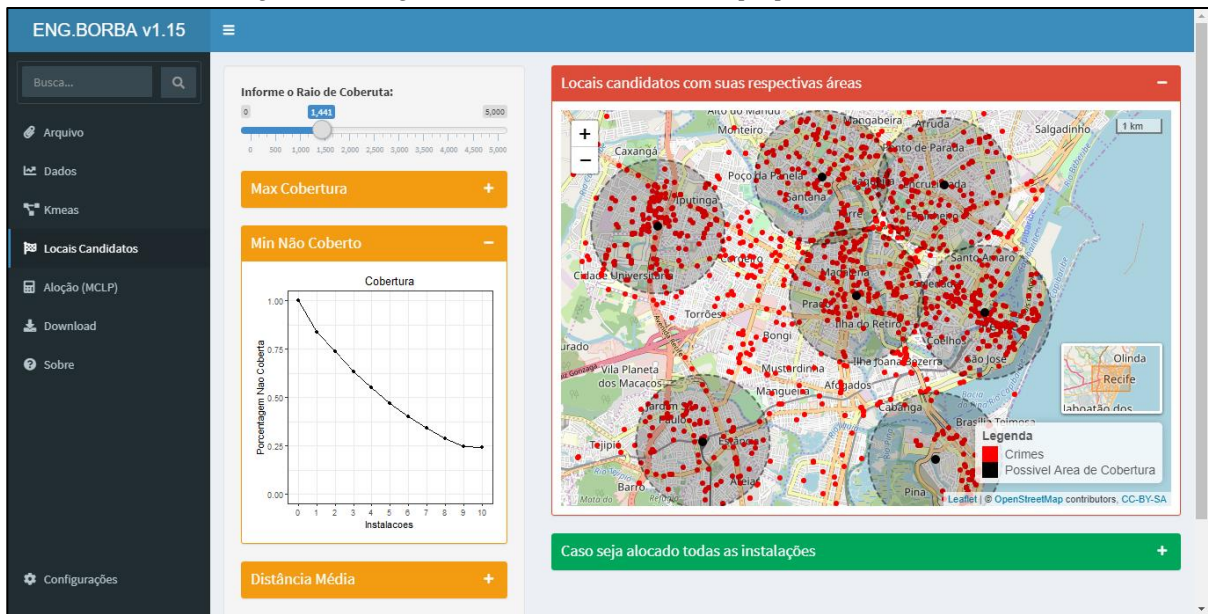
Em primeiro lugar, a barra corrediça graduada “Informe o Raio de Cobertura” é utilizada pelo usuário para informar o raio de atuação que as instalações irão cobrir quando forem alocadas. Portanto, é permitido ao usuário explorar com precisão o comportamento de diversos parâmetros à medida que o raio é alterado automaticamente no SIG.

Esses parâmetros são apresentados dentro das caixas abaixo da barra corrediça graduada. Que ao clicar no sinal de adição situado no lado superior direito da caixa, é aberto o gráfico correspondente a análise solicitada. Nesse sentido, a primeira caixa é a de “Max Cobertura” que pode ser visualizada na Figura 18, informando o ganho que é obtido na cobertura à medida que é aumentado o número de novas instalações alocadas, chegando até o limite de locais candidatos anteriormente definidos pelo usuário na página Kmeans.

Uma análise semelhante a essa, é obtida através da segunda caixa “Min Não Coberto” apresentada na Figura 19. Na caixa “Min Não Coberto” é demonstrado a análise correspondente a porcentagem de crimes que vão deixando de estar descobertos à medida que mais instalações vão sendo alocadas.

Já na terceira caixa “Distância Média” apresentada na Figura 20, tem sua análise voltada a calcular a redução da quilometragem média entre as instalações e os crimes cometidos, à medida que novas instalações vão sendo alocadas. E para essa complementar informação a quarta caixa “Desvio Médio” rapidamente demonstrada na Figura 17, informa o desvio padrão médio da quilometragem entre os crimes e as instalações.

Figura 19 – Página Locais Candidatos do SIG proposto, não coberto



Fonte: O Autor (2019)

Nesse sentido, o conjunto dessas quatro análises simultâneas, permite ao usuário obter uma ampla compreensão do problema que está sendo enfrentado. Fornecendo assim, um melhor direcionamento em sua tomada de decisão para estabelecer a quantidade de instalações a ser alocada e área a ser coberta.

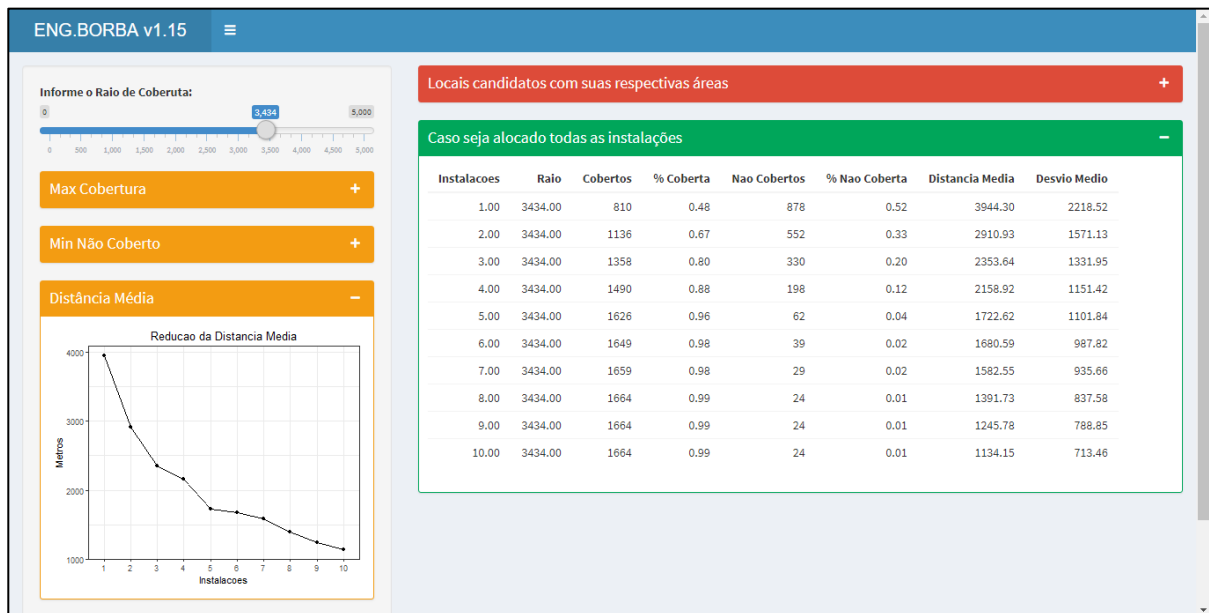
Para facilitar ainda mais essa tomada de decisão para o usuário, o ambiente de trabalho da página Locais Candidatos, é dividido em dois segmentos de caixas, são chamados de “Locais candidatos com suas respectivas áreas” e “Caso seja alocado todas as instalações” apresentadas nas Figuras 17, 18, 19 e 20.

No segmento “Locais candidatos com suas respectivas áreas” inicialmente apresentado na Figura 17, é observado que os locais candidatos nesse mapa, possuem a georreferência da área de atuação das instalações. Facilitando a interpretação do problema para o usuário, que deseja ter um conhecimento visual dos limites de cobertura de cada instalação.

Nesse sentido, assim que o usuário define um raio de cobertura no segmento “Informe o Raio de Cobertura”, automaticamente os limites de cobertura são atualizados no mapa. Essa mudança pode ser percebida ao analisar a diferença de área entre a Figura 17 e a Figura 19 que foram atribuídos diferentes valores para o raio. E da mesma forma que no “Mapa Crimes”, nesse mapa também consta uma interface interativa de movimentação, possibilitando ao usuário navegar por todas as direções, incluindo as funções de zoom. Já a função de *pop-up* nesse mapa, é habilitada ao clicar no ponto preto ou na circunferência na sua proximidade (Figura 17 e

Figura 18), informando que aquele ponto é uma possível instalação a ser alocada e sua respectiva área de cobertura.

Figura 20 – Página Locais Candidatos do SIG proposto, relatório



Fonte: O Autor (2019)

Já no segmento “Caso sejam alocadas todas as instalações” apresentado na Figura 20, é demonstrado para o usuário um relatório contendo todas as análises provenientes do processo de localização das instalações. Logo, esse relatório tem por objetivo demonstrar de forma numérica, a sensibilidade na qual a todos os parâmetros sofrem ao adicionar uma nova instalação até atingir o número máximo de locais candidatos.

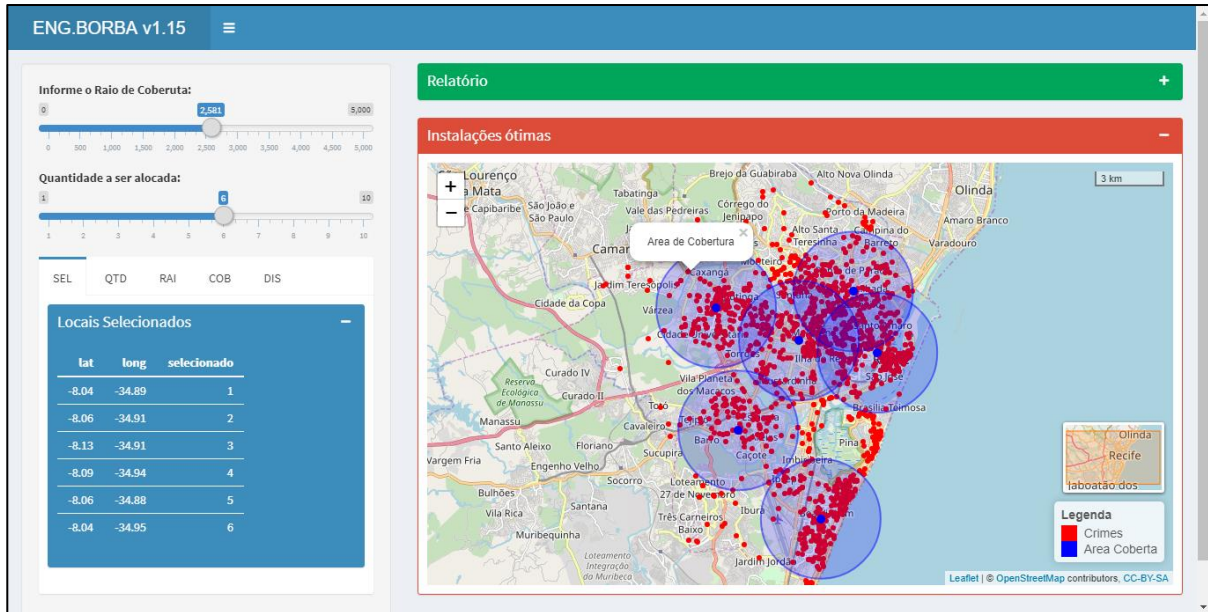
Portanto, o usuário na página de Locais Candidatos consegue manejar uma grande diversidade de análises. Possibilitando que o usuário desenvolva a sua diretriz no momento de estabelecer os principais parâmetros do problema de *Maximal Covering Location Problem* (MCLP).

Seguindo com as análises do design do SIG proposto, na Figura 21 é apresentado a página Alocação (MCLP). E é observado que a tela é dividida na barra branca e no ambiente de trabalho, constatando que a barra cinza foi desativada para maximizar a área de interação do usuário para com o SIG.

Dessa forma, é apresentado que a barra branca da Figura 21 possui três segmentos, no qual, dois deles são barras corrediças, chamadas de “Informe o Raio de Cobertura:” e “Quantidade a ser alocada:”. Que funcionam exatamente como as barras corrediças supracitadas, acrescentando que na barra “Quantidade a ser alocada:”, o usuário define

exatamente a quantidade que instalações a ser alocada entre as possibilidades existentes, definidas na página Kmeans do SIG.

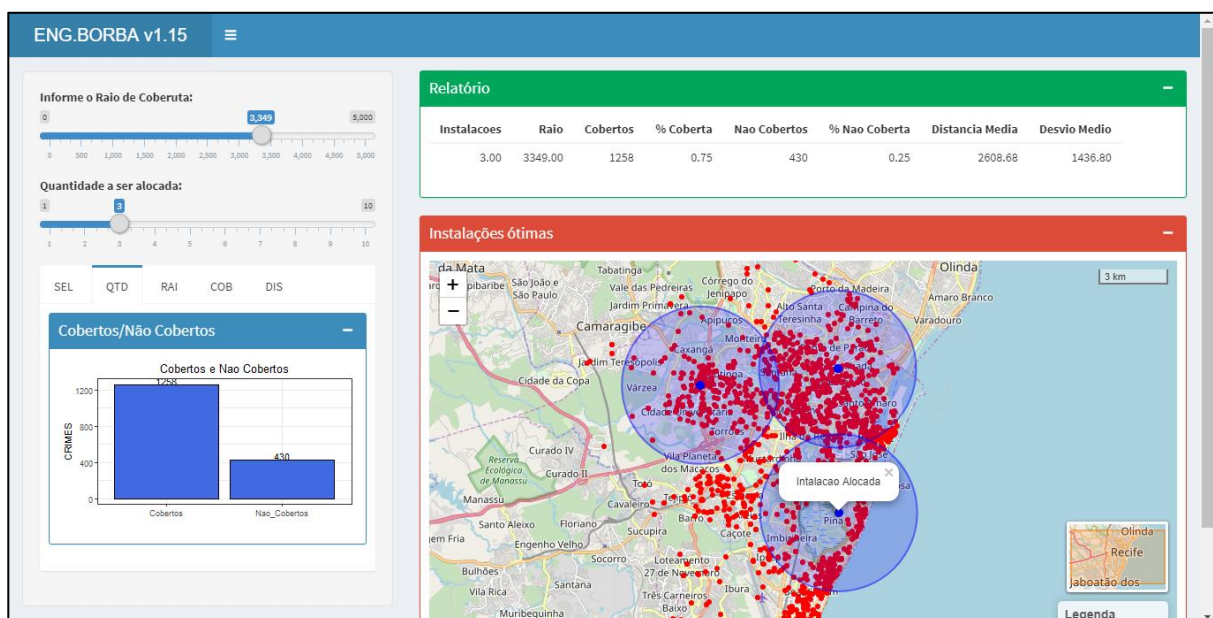
Figura 21 – Página Alocação (MCLP) do SIG proposto, quantidade



Fonte: O Autor (2019)

Já no terceiro segmento da barra branca na Figura 21, é apresentado uma caixa contendo cinco guias, chamadas de “SEL”, “QTD”, “RAI”, “COB” e “DIS”. Cada uma dessas guias é responsável por informar ao usuário uma análise diferente, que juntas foram obtidas durante o processo de alocação definitiva das instalações.

Figura 22 – Página Alocação (MCLP) do SIG proposto, locais selecionados

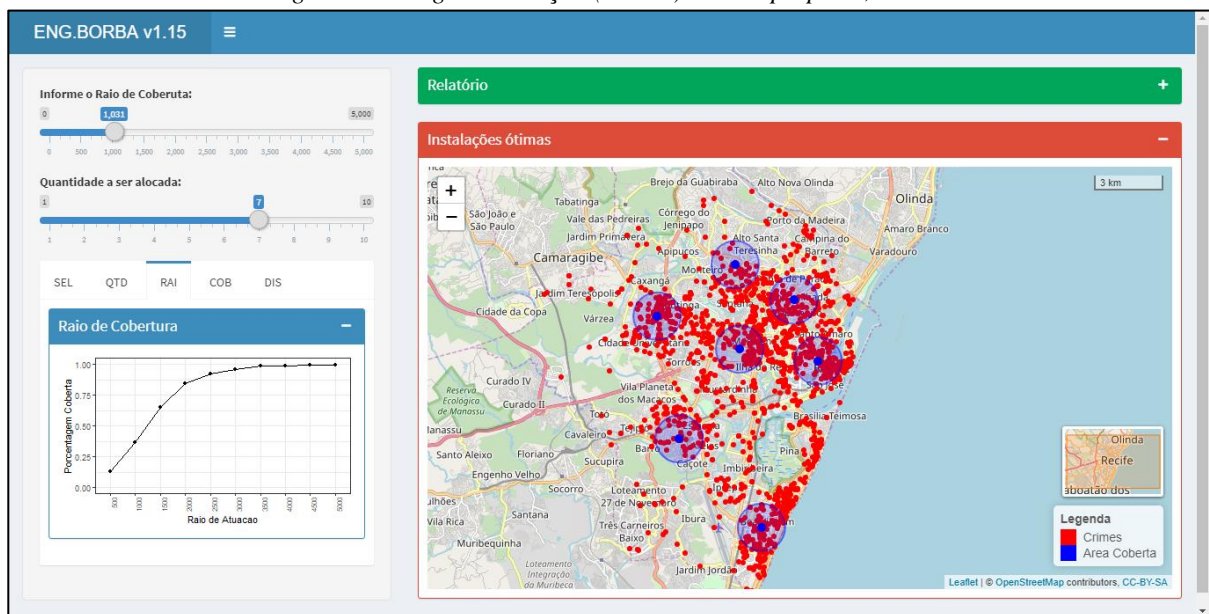


Fonte: O Autor (2019)

Nesse sentido, a guia “SEL”, é responsável por apresentar em forma de tabela quais são as coordenadas que foram selecionadas para instalar uma unidade policial dentre as coordenadas dos locais candidatos. Fazendo uso das informações fornecidas pelo usuário com relação ao raio e a quantidade de instalações a ser alocadas para maximizar a quantidade de crimes que serão cobertos. Portanto, na guia “QTD”, é apresentado um gráfico de barras (Figura 22) que é responsável por ilustrar a quantidade de crimes que estão cobertos e descobertos após a alocação das instalações.

Já na guia “RAI” apresentada a Figura 23, o objetivo é apresentar para o usuário uma análise de sensibilidade com relação ao raio de cobertura. Ou seja, é incrementado o tamanho do raio de forma gradual, permanecendo fixa a quantidade de instalações. Dessa forma, é apresentado em que momento é possível cobrir 100% dos crimes, ou mesmo a quantidade que o usuário desejar.

Figura 23 – Página Alocação (MCLP) do SIG proposto, raio



Fonte: O Autor (2019)

As análises fornecidas pelas guias “COB” e “DIS” na página Alocação (MCLP), são as mesmas análises fornecidas pelas guias “Max Cobertura” e “Distância Média” respectivamente da página Locais Candidatos. A única diferença que ocorre nesse novo cenário é que, as análises da página Alocação (MCLP) estão apenas para os locais ótimos selecionados e não para todos os possíveis como é realizado na página Locais Candidatos. E da mesma forma que na página Locais Candidatos, a página Alocação (MCLP) tem seu ambiente de trabalho dividido em dois segmentos de caixa, chamados de “Relatório” e “Instalações Ótimas”, como pode ser visto na Figura 21, Figura 22 e Figura 23.

O mapa fornecido pela caixa “Instalações Ótimas” possui pequenas modificações com relação aos mapas anteriores. Essas pequenas modificações podem ser observadas através das cores e dos *pop-ups*. É o caso da cor azul que representa a área coberta pela instalação alocada em seu centro. Já a caixa “Relatório” apresentado na Figura 22, fornece um rápido e claro resultado da alocação das instalações, com informações numéricas sobre a quantidade de instalações, a cobertura, a distância média, os crimes não cobertos e diversas porcentagens com relação a esses dados.

Porém, para que todo sistema de informação seja bem estruturado, é necessário que o mesmo seja capaz de coletar as informações do usuário, processe todas essas informações e permita ao usuário a coleta desses resultados. Nesse sentido a página Download do SIG proposto, apresentado na Figura 24, tem por objeto permitir ao usuário o *download* de qualquer um dos resultados desenvolvidos pelo SIG.

Nesse sentido, a página Download do SIG é dividida em uma barra branca e o seu ambiente de trabalho. No qual, a barra branca contém dois segmentos. O primeiro é uma caixa de seleção onde o usuário pode utilizar para selecionar qualquer relatório que desejar obter do SIG. Que entre eles estão o relatório de locais candidatos, locais selecionados, resultado final e sensibilidade.

Figura 24 – Página Download do SIG proposto

The screenshot shows the 'ENG.BORBA v1.15' interface. On the left is a navigation menu with options: Arquivo, Dados, Kmeas, Locais Candidatos, Alocação (MCLP), Download, Sobre, and Configurações. The main area has a search bar and a dropdown menu titled 'Escolha o relatório que deseja baixar' with options: Nenhum, Locais Candidatos, Locais Selecionados, Resultado Final, and Sensibilidade. To the right is a table titled 'RELATÓRIO PARA DOWNLOAD' with the following data:

Instalacoes	Raio	Cobertos	% Coberta	Nao Cobertos	% Nao Coberta	Distancia Média	Desvio Medio
1.00	1969.00	347	0.21	1341	0.79	4672.32	2750.41
2.00	1969.00	603	0.36	1085	0.64	3509.59	2581.92
3.00	1969.00	814	0.48	874	0.52	2979.76	2530.43
4.00	1969.00	990	0.59	698	0.41	1921.64	1286.84
5.00	1969.00	1125	0.67	563	0.33	1732.36	1270.20
6.00	1969.00	1255	0.74	433	0.26	1479.02	1027.79
7.00	1969.00	1346	0.80	342	0.20	1390.98	984.94
8.00	1969.00	1434	0.85	254	0.15	1277.93	922.17
9.00	1969.00	1513	0.90	175	0.10	1165.74	847.29
10.00	1969.00	1521	0.90	167	0.10	1134.15	713.46

At the bottom of the main area is a 'Download' button.

Fonte: O Autor (2019)

Já o segundo segmento, é um botão de *download*, que após a seleção do relatório desejado, o botão é habilitado para criar e efetuar a transferência do arquivo para o computador

do usuário. O botão ainda permite que o usuário escolha qual diretório de seu computador o SIG pode salvar o relatório selecionado.

Já o ambiente de trabalho da página Download apresentado na Figura 24, tem como objetivo através de seu único segmento, permite ao usuário uma visão prévia do relatório que vai ser baixado em seu computador. Dessa forma o usuário pode revisar qualquer informação antes do relatório ser criado, fornecendo assim, rapidez e eficiência caso seja preciso alterar qualquer informação fornecida pelo SIG.

É possível analisar que na Figura 24 na barra cinza que ainda faltam duas páginas a serem apresentadas, chamadas de Sobre e Configurações. A página Sobre tem por objetivo apresentar informações do desenvolvimento do SIG. Informações essas, relacionadas a versão do SIG, pessoas envolvidas em sua criação, data da última atualização, funções em desenvolvimento e descrições dos tipos de problema que o SIG é capaz de solucionar.

A página Configurações do SIG proposto, encontra-se atualmente em fase de desenvolvimento. Porém seu escopo já está definido para ser responsável por alterar diversos componentes internos do SIG. Esses componentes vão desde configurações visuais a modificações internas do algoritmo base, que serão realizadas sem a necessidade de conhecimentos específicos.

Por fim, torna-se notório a facilidade que o usuário tem ao interagir com o SIG proposto para solucionar a problemática dessa pesquisa. As caixas de análises, de seleções, de mapas e os botões integrados no design do SIG, permite que qualquer pessoa seja capaz de manuseá-lo. E dessa forma, acaba se tornando mais fácil de desenvolver o estudo de caso dessa pesquisa, que consta no capítulo seguinte.

6 ESTUDO DE CASO

O estudo realizado nesse trabalho foi aplicado em um caso real onde o problema de localização de instalações policiais é fundamental para a redução dos índices criminais e a melhoria do atendimento da demanda existente. Dessa forma, ao longo desse capítulo será aplicado o Sistema de Informação Gerencial desenvolvido no capítulo 5 e apresentado com base no mesmo, a região onde foi realizada esta pesquisa, a fonte dos dados coletados, o processo de seleção dos locais candidatos, a alocação otimizada das instalações e uma comparação do resultado ótimo obtido com o cenário atual vivenciado na região.

6.1 Local de estudo

O local escolhido para o desenvolvimento dessa pesquisa foi a cidade do Recife, município brasileiro e capital do estado de Pernambuco, que é situado na Região Nordeste do país. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), a cidade possui uma área territorial de 218,843 km² e com base no último censo de 2010, é estimado que para 2018 a cidade possui aproximadamente 1.637.834 habitantes, constatando de uma densidade populacional de 7.484,06 hab/km².

Nesse sentido, quando se trata de segurança para essa massa populacional do Recife, a SDS de Pernambuco integra as ações do governo com vistas a preservar a ordem pública e a segurança das pessoas e bens dentro do estado, sendo responsável por definir onde as unidades policiais serão alocadas (GUSMÃO *et al.*, 2019). Porém, mesmo o trabalho eficaz desenvolvido pela SDS de Pernambuco, os índices criminais na cidade do Recife vem aumentando nos últimos anos (12º ANUÁRIO DE SEGURANÇA PÚBLICA, 2018). Alertando a necessidade de mais estudos na região, com o objetivo de desenvolver novas técnicas para a redução desses índices criminais.

E para que novos estudos e técnicas sejam desenvolvidas pelas universidades, é preciso definir uma boa base de dados para executar a pesquisa. Nesse sentido, a próxima seção visa apresentar a base de dados que foi utilizada para o desenvolvimento dessa pesquisa.

6.2 Base de dados

A base de dados utilizada nessa pesquisa foi obtida a partir da plataforma “Onde Fui Roubado”, que se trata de uma plataforma colaborativa contra a violência. Na plataforma, os usuários em seu anonimato marcam em um mapa, os locais em que sofreram furtos, roubos, sequestro relâmpago, assalto a residências, lojas e outros tipos de crimes.

A plataforma foi desenvolvida inicialmente em 2012 por dois estudantes (Filipe Norton e Márcio Vicente) do curso de Ciências da Computação da Universidade Federal da Bahia (UFBA). Rapidamente a plataforma foi reconhecida e esteve operando em todo o Brasil, apresentando os pontos em que as pessoas foram vítimas de algum tipo de crime. Porém a plataforma veio a ficar *offline* em 2019.

A plataforma era de grande ajuda para a população, pois facilitava o acesso ao conhecimento das regiões mais violentas de uma determinada cidade. No qual, através de etiquetas marcadas no mapa com cores diferentes, possibilitava a população distinguir o tipo de crime e analisa-lo em maiores detalhes cada um deles.

Antes da plataforma ser desativada, vários estudos já vinham sendo desenvolvidos na área de segurança pública, através da base de dados fornecida livremente pela plataforma “Onde fui Roubado”. Nesse sentido, a Universidade Federal de Pernambuco (UFPE), mais especificamente o Grupo de Pesquisa em Sistemas de Informação e Decisão (GPSID) com a posse desses dados, permitiu o desenvolvimento desta pesquisa.

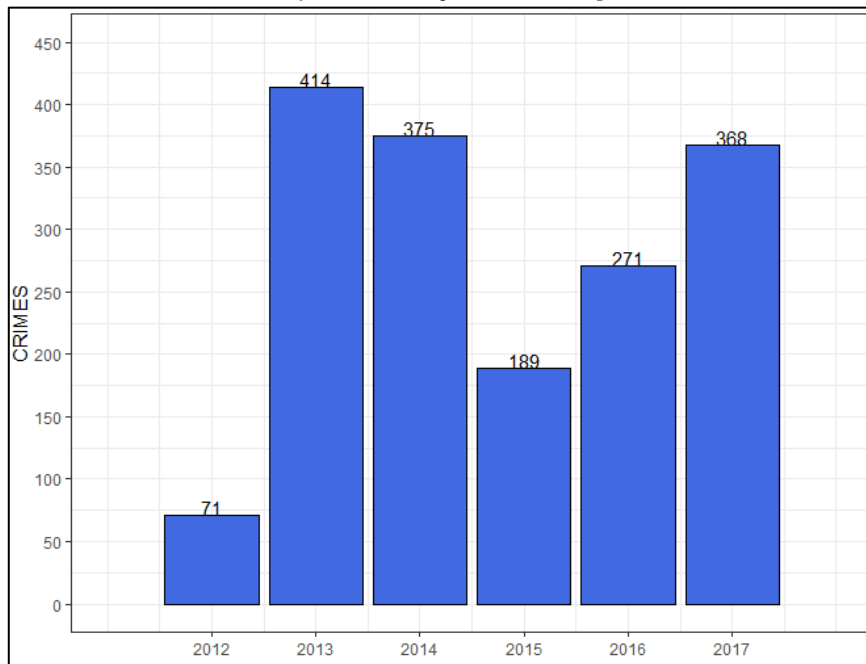
6.3 Dados Iniciais

Para o desenvolvimento desta pesquisa, foram coletados 1.688 dados de ocorrências georreferenciados na cidade do Recife, a partir da plataforma colaborativa “Onde Fui Roubado”. Esses dados são correspondentes ao intervalo de observação que se inicia em janeiro de 2012 e vai até julho 2017.

Nesse sentido, no Gráfico 8 é apresentado a relação dos crimes que foram cometidos por ano, referente a essa base de dados. Como os dados foram preenchidos por qualquer pessoa (geralmente a vítima), por se tratar de uma plataforma colaborativa, é possível observar que no ano de 2012 houve um menor número de ocorrências, ao ser comparado com os demais anos.

Uma explicação plausível para essa divergência do ano de 2012, é a popularidade da plataforma que tem seu início no mesmo ano. Logo, muitas pessoas ainda não conheciam a ferramenta e diversos dados não foram computados. Dessa forma, se torna notório que diversos fatores influenciam no somatório total de crimes que ocorrerão por ano. Porém, esses fatos não descartam a utilidade da plataforma e o uso de seu banco de dados para o desenvolvimento de novos estudos.

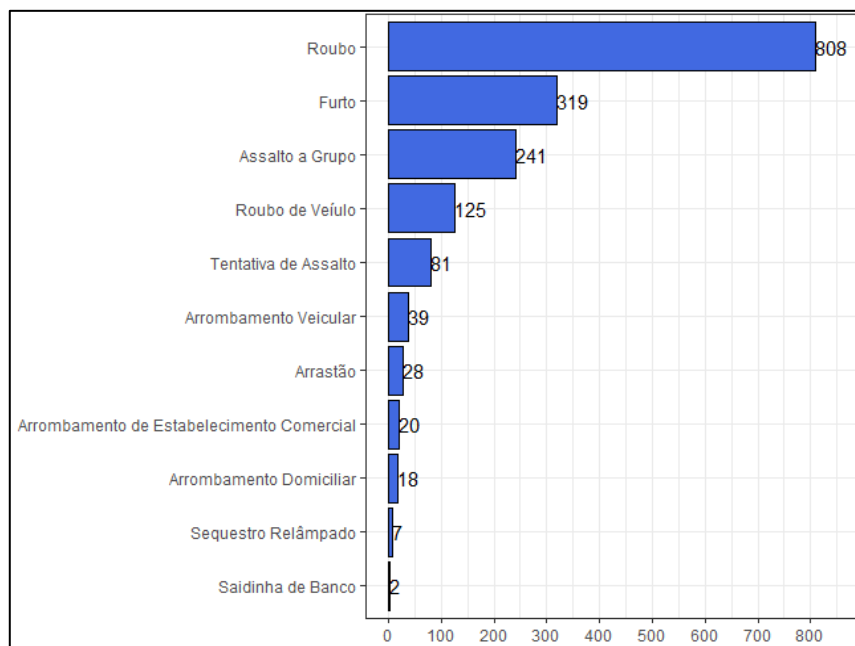
Gráfico 8 – Relação de Crimes por Ano



Fonte: O Autor (2019)

Seguindo com a apresentação estatística dos resultados iniciais obtidos através do banco de dados. O Gráfico 9 demonstra o quantitativo de crimes ocorridos por classificação, ou seja, os tipos de crimes mais recorrentes cometidos durante esse espaço de tempo observado.

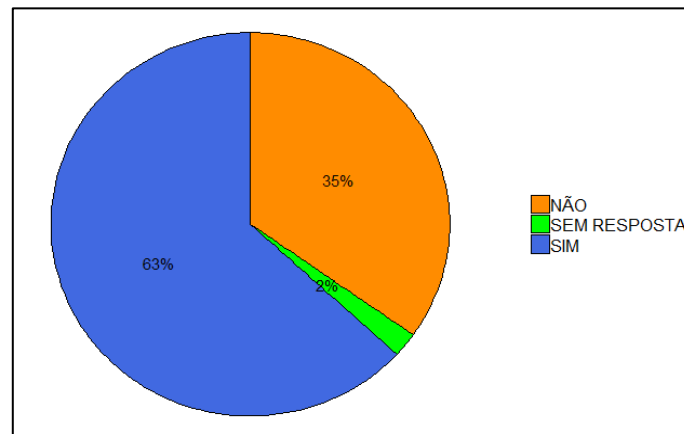
Gráfico 9 – Quantitativo de Crimes por Tipo



Fonte: O Autor (2019)

No Gráfico 10, é apresentado o percentual de vítimas que se preocuparam em registrar o boletim de ocorrência policial, seja ele, efetuado via internet ou por uma unidade policial mais próxima. Nesse sentido, o boletim de ocorrência policial, que também é chamado de BO, trata-se de um registro oficial e detalhado do ato criminoso ocorrido, servindo de extrema relevância para as autoridades encarregadas da segurança pública.

Gráfico 10 – Percentual que Registrou Boletim de Ocorrência

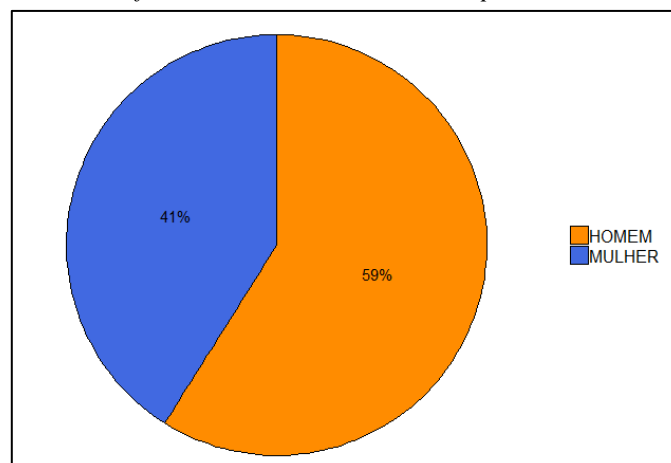


Fonte: O Autor (2019)

Com essas informações oficiais, as autoridades conseguem desenvolver um mapa de crimes que identifica as diversas concentrações desses eventos. Facilitando para os tomadores de decisão, a distribuição correta dos recursos operacionais para garantir uma maior eficiência.

E o que é apresentado na base de dados desse estudo, através do Gráfico 10. É que uma grande porcentagem das vítimas, representando 35% das observações, não realizam o registro oficial do boletim de ocorrência. O que pode inviabilizar a elaboração adequada do mapa de crime e, conseqüentemente, comprometer a distribuição correta dos recursos operacionais da polícia.

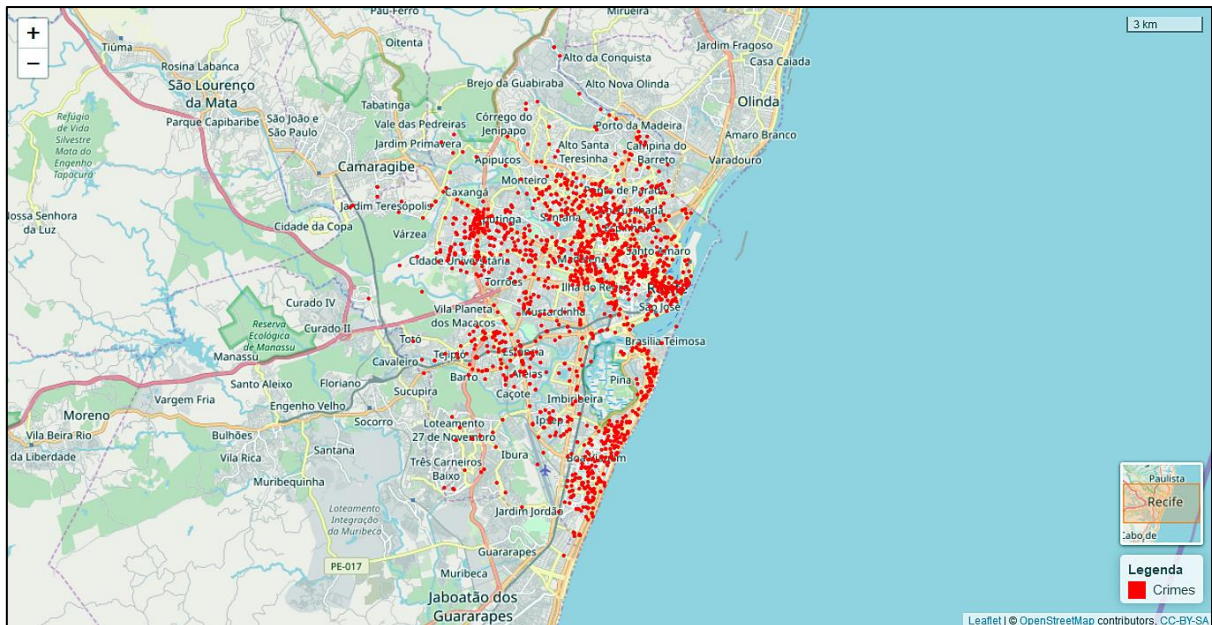
Gráfico 11 – Percentual de Vítimas por Gênero



Fonte: O Autor (2019)

A base de dados da plataforma “Onde Fui Roubado” também contém informações sobre gênero, demonstrando assim, um quantitativo de homens e mulheres que foram vítimas de algum tipo de crime. Nesse sentido, o Gráfico 11 tem como objetivo apresentar em forma de gráfico de setores, a porcentagem de homens e mulheres que foram vítimas, o que deixa claro que os homens estão na frente com 59% dos casos ocorridos.

Mapa 1 – Mapa de Crimes em Recife



Fonte: O Autor (2019)

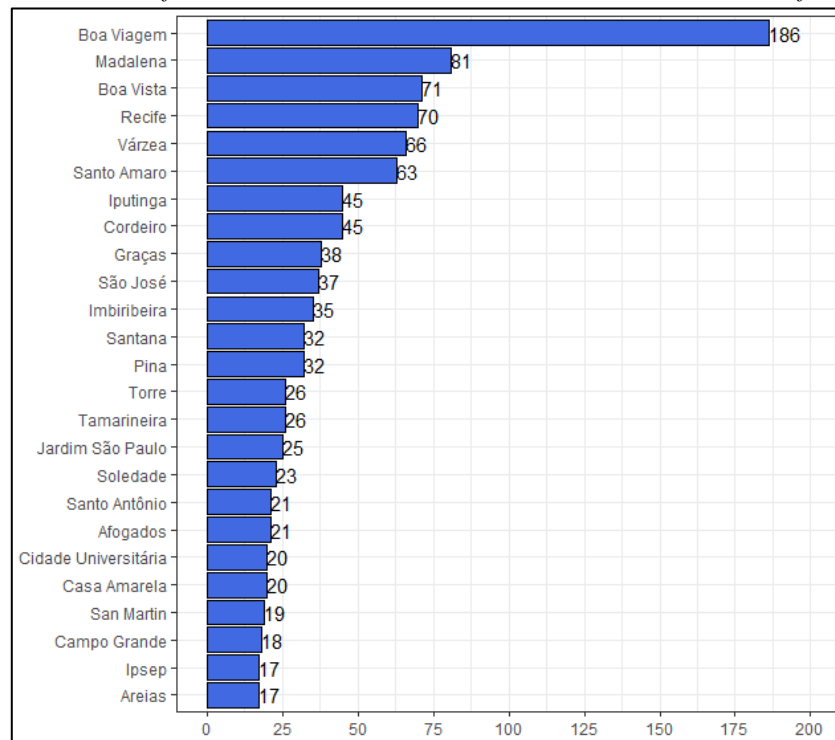
Nesse sentido, com todas essas informações supracitadas, foi utilizado o SIG apresentado no capítulo 5, para coletar as informações da base de dados e criar um mapa de crimes da cidade do Recife. Como foi afirmado, o mapa de crimes vai permitir aos tomadores de decisão obter uma visão geral dos crimes ocorridos durante o intervalo de tempo observado. Logo, no Mapa 1 é apresentado em pontos vermelhos os crimes ocorridos durante o intervalo de 2012 a 2017 na cidade do Recife.

Dessa forma, com o desenvolvimento do mapa de crimes da cidade do Recife, se torna notório a visualização de que em alguns bairros existe uma concentração de crimes bem maior do que outros. Portando no Gráfico 12 é apresentado um gráfico de barras com os bairros da cidade do Recife que mais tiveram ocorrências de crimes.

Nesse sentido, em primeiro lugar em números de ocorrências, se encontra o bairro de Boa Viagem, chegando a abranger mais que o dobro de ocorrências do bairro que se encontra em segundo lugar, Madalena. Essas informações destacam a importância do mapa de crimes para

os tomadores de decisão, que precisam planejar bem os locais que devem alocar as unidades policiais, para aumentar a eficiência.

Gráfico 12 – Bairros Com Mais Ocorrência de Crimes em Recife



Fonte: O Autor (2019)

Com a utilização do SIG proposto, foi possível realizar e extrair esse estudo inicial e estatístico, que fornece dados relevantes para o tomador de decisão. Logo, a partir desses estudos iniciais, é possível fazer uso do SIG proposto para encontrar os locais candidatos que melhor atendem as necessidades dessa demanda existente.

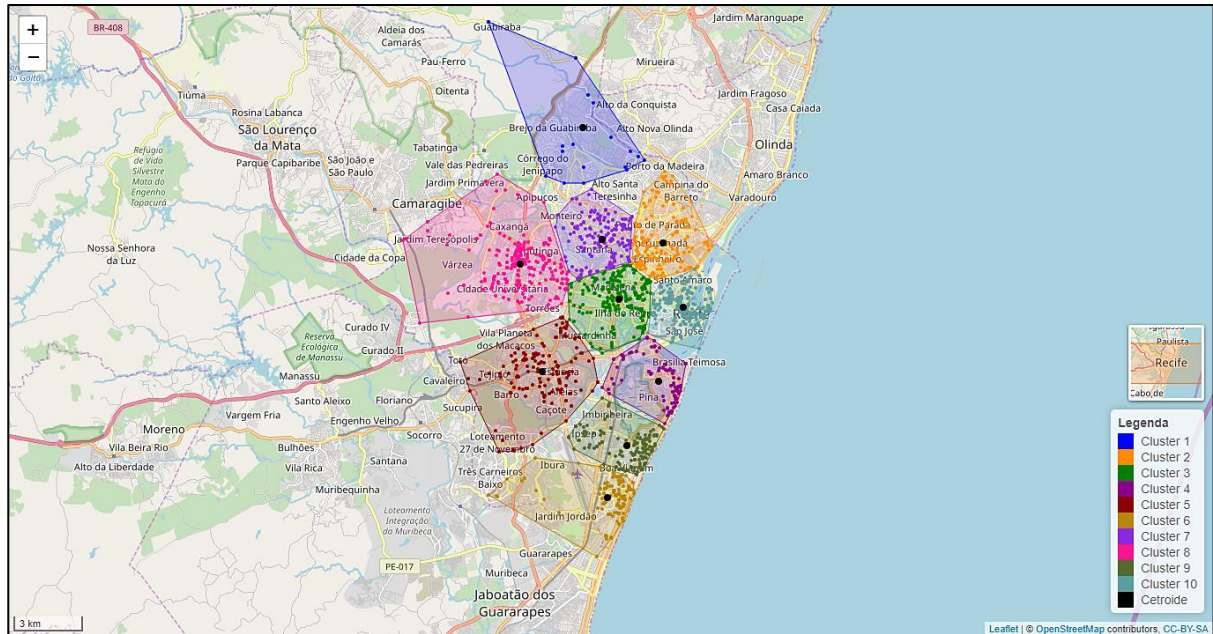
6.4 Locais Candidatos

Antes de efetuar a localização exata e otimizada das instalações policiais no território da cidade do Recife, é preciso realizar um estudo inicial para definir quais serão os locais candidatos para se alocar as instalações. Pois, nesses locais a polícia militar (PM) desempenha suas funções e a cada turno (madrugada, manhã, tarde ou noite) os veículos deixam essas instalações para patrulhar uma determinada área de cobertura.

Nesse sentido, é necessário que as instalações da PM estejam localizadas em pontos que minimizem o tempo de resposta do veículo às ocorrências. Portanto, foi utilizado o SIG proposto e a base de dados em estudo, para agrupar os crimes em clusters, mediante a semelhança da região onde os crimes estão ocorrendo com maior frequência. Logo, no Mapa 2,

é ilustrado as dez regiões em que os crimes estão ocorrendo com maior frequência e semelhança na cidade do Recife.

Mapa 2 – Clusters de Crimes na Cidade do Recife



Fonte: O Autor (2019)

Ainda com o SIG proposto, é encontrado o centroide de cada uma das regiões que são delimitadas pelos clusters. Sendo cada centroide, um local candidato para se alocar uma instalação policial. Dessa forma, no Quadro 2 é apresentado esses centroides em forma de coordenadas geográficas, deixando claro, todos os dez locais candidatos.

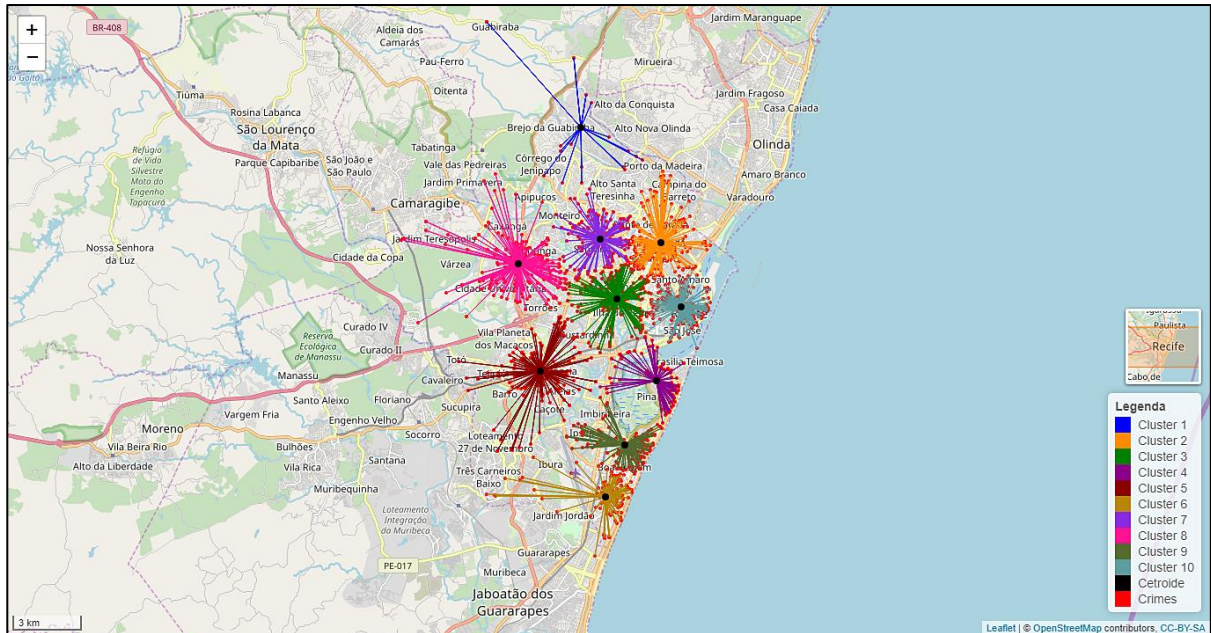
Quadro 2– Coordenadas dos locais candidatos

Locais Candidatos	Latitude	Longitude
Candidato 1	-7.992144	-34.92050
Candidato 2	-8.036942	-34.88901
Candidato 3	-8.058401	-34.90614
Candidato 4	-8.090015	-34.89074
Candidato 5	-8.086555	-34.93611
Candidato 6	-8.135032	-34.91076
Candidato 7	-8.035501	-34.91289
Candidato 8	-8.045003	-34.94490
Candidato 9	-8.115055	-34.90337
Candidato 10	-8.061584	-34.88112

Fonte: O Autor (2019)

Com o Mapa 3 se torna notório que a cobertura total da demanda de cada cluster, depende do raio de atuação da instalação. Ou seja, quanto maior o raio, maior a quantidade de crimes que serão cobertos. Porém, vale salientar que quanto maior o raio, maior será o tempo que a viatura irá levar para sair da instalação e chegar no local da ocorrência.

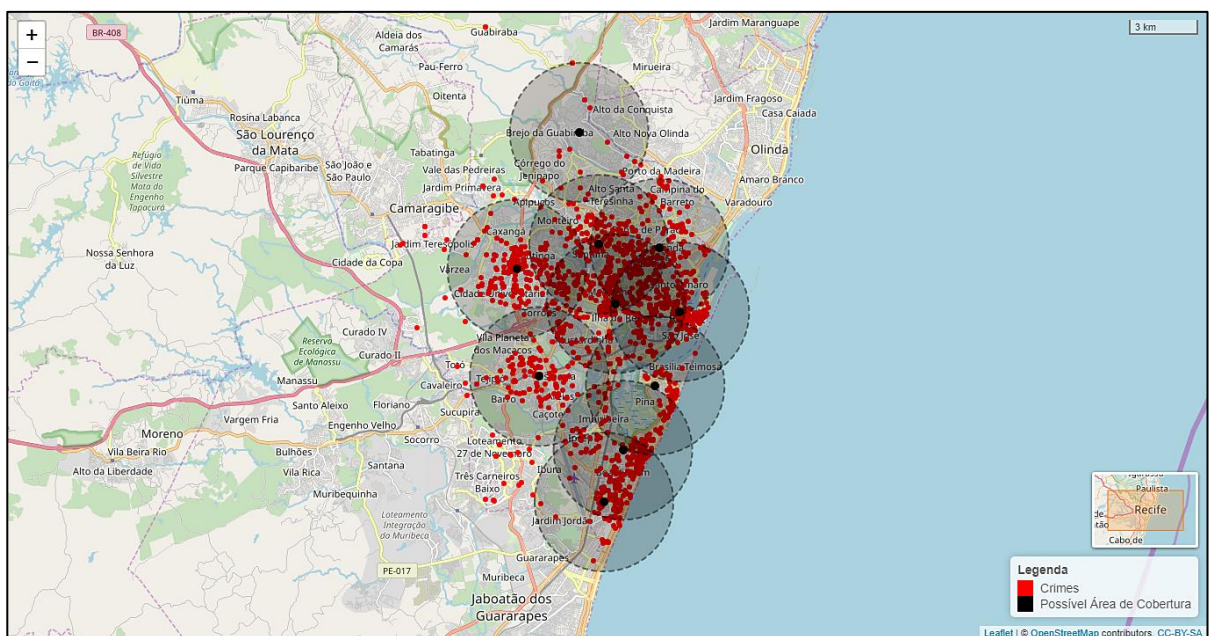
Mapa 3 – Cluster Distância



Fonte: O Autor (2019)

Porém, como supracitado no capítulo 2, diversos fatores influenciam no tempo de resposta e não há um valor bem definido para uma distância de serviço ou um tempo de resposta. É de responsabilidade do decisor definir esse parâmetro de acordo as metas de segurança da região.

Mapa 4 – Georreferência dos Locais Candidatos



Fonte: O Autor (2019)

Nesse sentido, o Mapa 4 é ilustrado os dez locais candidatos com um raio de cobertura de três quilômetros. Deixando claro como será desenvolvido a análise dessa pesquisa com relação aos crimes que serão cobertos por cada raio de cobertura a ser testado.

Logo, como não é fornecido um raio específico e exato de cobertura para as instalações da cidade do Recife. Esta pesquisa irá explorar na próxima seção, diversos cenários entre os valores de um a cinco quilômetros de cobertura. Para análise a sensibilidade desse parâmetro e verificar o comportamento da alocação otimizada dessas instalações para cada raio de cobertura.

6.5 Alocação ótima

Após a definição dos locais candidatos e a faixa de valores que será trabalhado com relação ao raio de cobertura. Se faz necessário a definição da restrição orçamentária, ou seja, a quantidade de instalações que serão alocadas.

Para isso, entre os dez locais candidatos, não foram selecionadas mais de cinco instalações. Essa escolha de apenas cinco instalações ocorreu devido ao interesse inicial em analisar o cenário ótimo obtido com o SIG proposto e as cinco instalações da PM do Recife que estão atualmente em funcionamento.

Quadro 3 – Cobertura ótima para cinco instalações alocadas

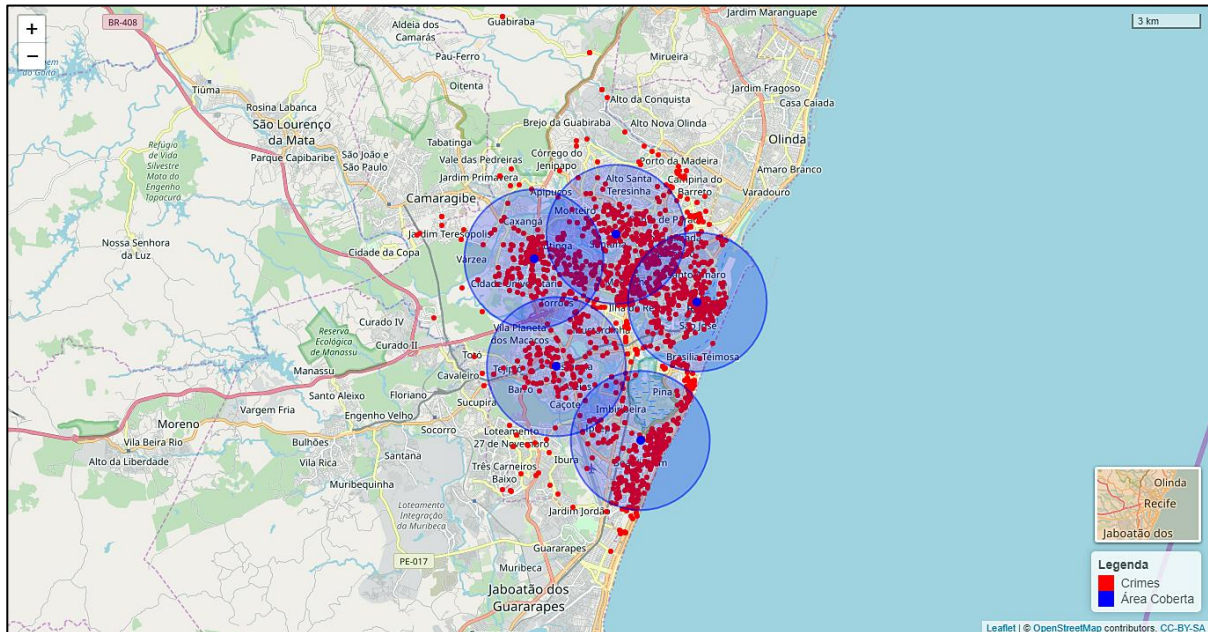
Raio de Cobertura (Km)	Números de Cobertos	Porcentagem Coberta	Representação Geográfica
1.00	582	34,5%	Apêndice A – Mapa 6
1.50	923	54,7%	Apêndice A – Mapa 7
2.00	1143	67,7%	Apêndice A – Mapa 8
2.50	1380	81,8%	Apêndice A – Mapa 9
3.00	1521	90,1%	Capítulo 6 – Mapa 5
3.50	1630	96,6%	Apêndice A – Mapa 10
4.00	1664	98,6%	Apêndice A – Mapa 11
4.50	1675	99,2%	Apêndice A – Mapa 12
5.00	1683	99,7%	Apêndice A – Mapa 13

Fonte: O Autor (2019)

Nesse sentido, foi utilizado o SIG proposto para realizar a alocação ótima das cinco instalações entre os valores de raio de 1.0 a 5.0 quilômetros. E dessa forma, foram obtidos os dados do Quadro 3, que apresenta o ganho de cobertura, através do aumento do raio de atuação

das cinco instalações alocadas. Deixado em evidência, que os valores do raio de 2.5, 3.0, e 3.5 quilômetros, proporcionam uma cobertura admissível das ocorrências, ou seja, 81,8%, 90,1% e 96,6%, respectivamente. No Mapa 5, são ilustradas as instalações ótimas alocadas, conforme o raio de cobertura de 3.0 quilômetros.

Mapa 5 – Instalações ótimas alocadas, com raio de 3.0 km de cobertura



Fonte: O Autor (2019)

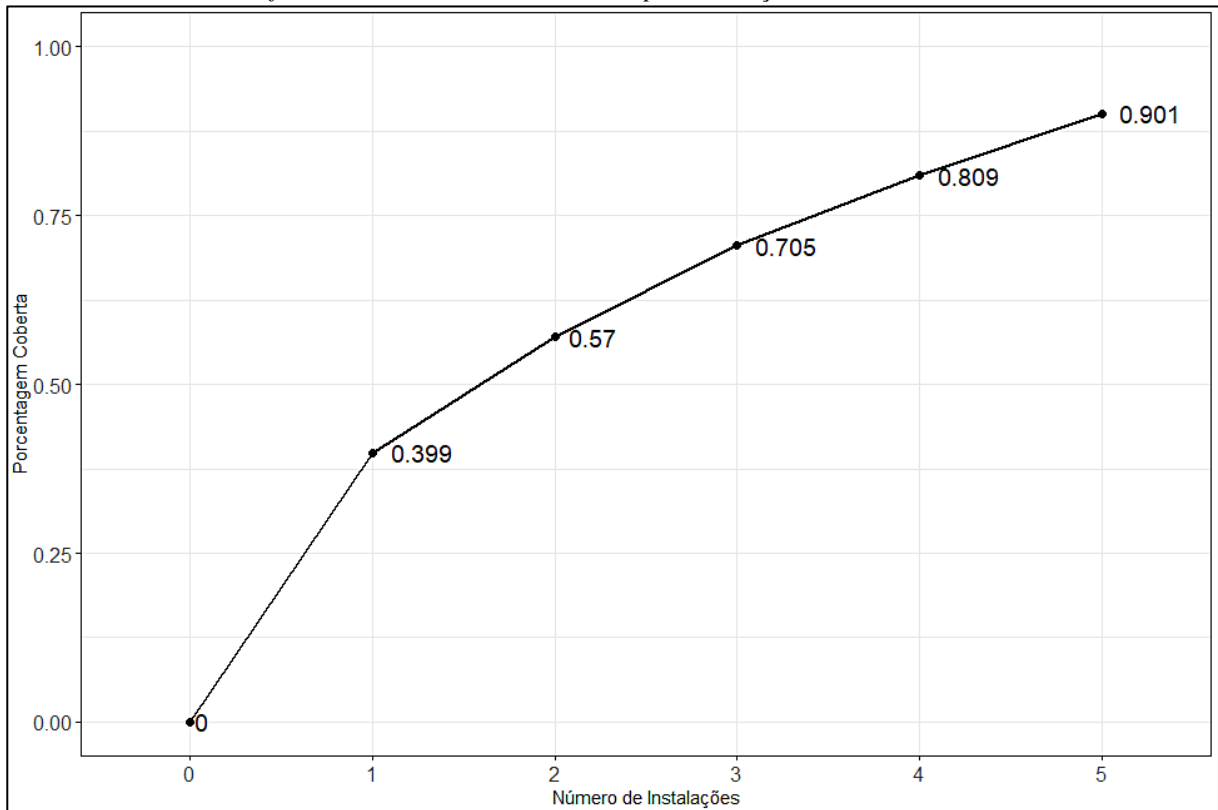
Portanto, na quarta coluna (Representação Geográfica) do Quadro 3, são apresentadas as referências das figuras para as instalações ótimas alocadas conforme o raio de cobertura avaliado. Ou seja, no Apêndice A deste trabalho, é possível observar a dinâmica realizada pelo SIG proposto no momento de localizar de forma ótima as instalações. Nesse sentido, conforme o raio de cobertura é fornecido, o SIG proposto vai buscar entre todos os locais candidatos, os locais que juntos possibilitam cobrir o maior número de ocorrências.

Essa afirmação, se torna clara ao observar os dados do Quadro 3 em conjunto com o Apêndice A. Onde a partir do primeiro incremento de 1.0 para 1.5 quilômetros do raio de cobertura, duas instalações mudaram de posição. E nos demais incrementos do raio, foram observados que em todos casos, ao menos uma das instalações ótimas mudou de posição. Sendo uma exceção apenas ao incremento de 2.5 para 3.0 quilômetros, onde as instalações ótimas permaneceram nas mesmas coordenadas. Provando assim, a eficiência do SIG proposto para localizar as instalações ao fazer uso do modelo *Maximal Covering Location Problem*.

Visando a obtenção de mais conhecimento através do SIG proposto, ao ser isolado um raio específico de cobertura, é possível analisar a sensibilidade com relação ao acréscimo de

cada instalação policial alocada. Nesse sentido, no Gráfico 13 é ilustrado o ganho de cobertura para cada instalação alocada com raio de três quilômetros, instalações essas, que foram representadas geograficamente pelo Mapa 5. Dessa forma, se torna mais fácil para o decisor analisar até que ponto é rentável o investimento em instalações, mediante a porcentagem ganha em demanda coberta por cada instalação.

Gráfico 13 – Incremento de cobertura por instalações, com raio de 3km



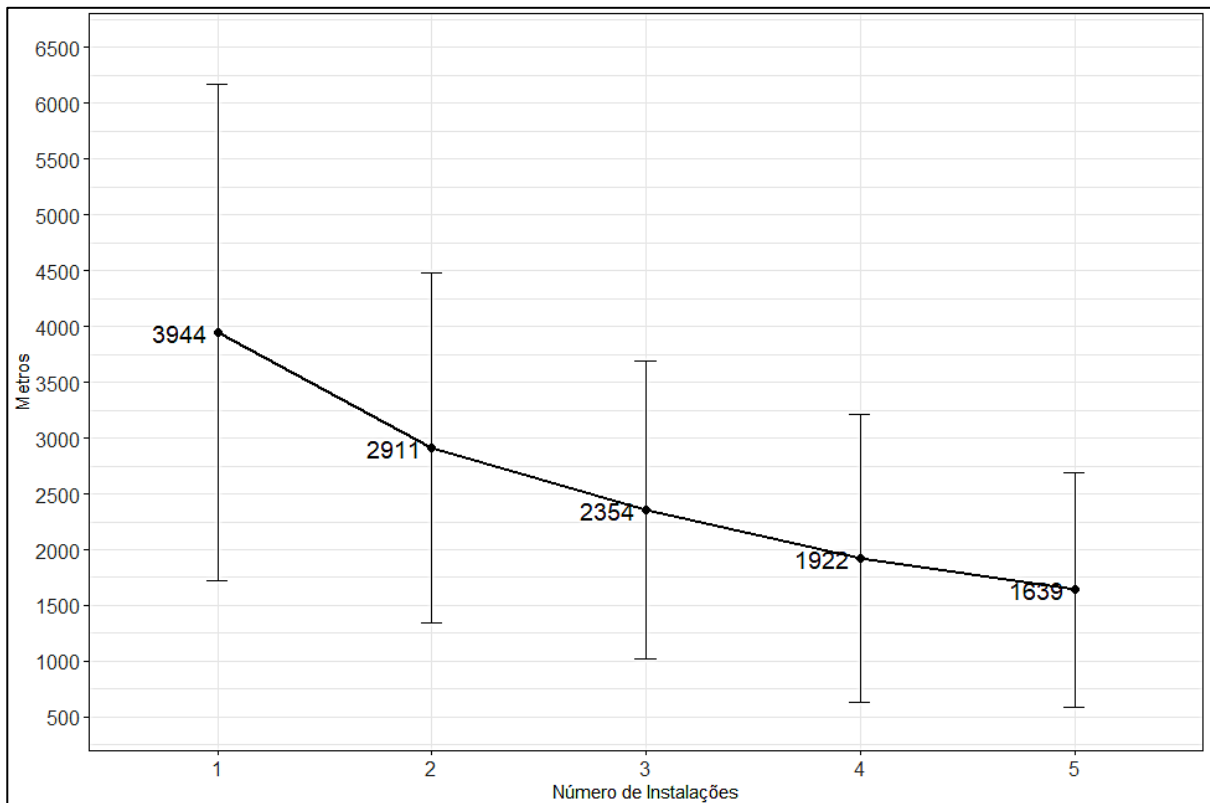
Fonte: O Autor (2019)

Portanto, torna-se notório que para o problema de localização de instalações envolvendo a máxima cobertura, que o problema envolve duas dimensões a ser analisada. No qual, a primeira dimensão seria o de fixar a quantidade de instalações e variar o raio de cobertura. E a segunda dimensão seria, o de fixar o raio de cobertura e variar a quantidade de instalações. Nesse sentido, o Apêndice B apresenta para esse problema, a análise de sensibilidade para essas duas dimensões em diversos cenários. Demonstrando assim, a quantidade de informações que são de extrema relevância para o decisor e o SIG proposto é capaz de apresentar.

Um outro aspecto de bastante importância para o tomador de decisão, é a análise da distância média para todas as ocorrências. Com essa informação o decisor pode verificar o quanto em quilometragem, as viaturas necessitam percorrer até conseguir chegar ao local da ocorrência.

Esse tipo de análise, implica não somente no tempo e no espaço, mas em análises financeiras que podem ser efetuadas com base nessas informações. Sejam elas voltadas para, as manutenções das viaturas ou para os gastos que são efetuados com combustíveis. Portanto, seguindo o exemplo de três quilômetros de raio, no Gráfico 14 é apresentado a análise da distância média percorrida em conjunto com os limites de confiança (desvio padrão médio) para se alcançar as ocorrências do banco de dados em estudo.

Gráfico 14 – Redução da distância média por instalações, com raio de 3km



Fonte: O autor (2019).

Em conclusão desta seção, é possível destacar que a alocação das instalações policiais no território da cidade do Recife foi plausível, mediante ao que foi projetado para executar o SIG proposto. Conseguindo bons resultados para pequenos raios de cobertura, como foi o exemplo de cinco instalações para três quilômetros, atingindo 90,1% das ocorrências e uma quilometragem média de 1,639 quilômetros.

E com base nessas informações supracitadas, a próxima seção tem por objetivo realizar um comparativo entre as informações obtidas pelo SIG proposto neste trabalho, e as instalações da PM do Recife que estão em funcionamento atualmente. No qual, em todos os casos, será usado um caso específico de raio de 3 quilômetros.

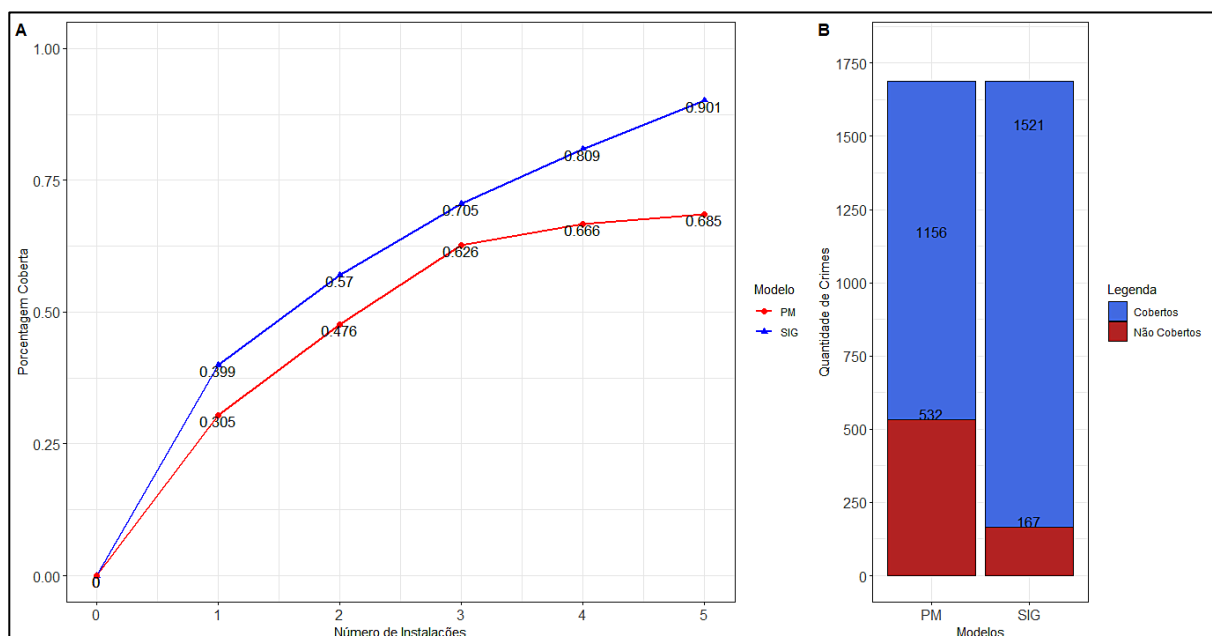
6.6 Comparação do cenário atualmente utilizado com o proposto

Após o desenvolvimento e aplicação teórica do SIG proposto na cidade do Recife, se teve o *insight* de comparar o cenário atual desempenhado com as localizações das unidades da PM (11° BPM -BATALHÃO 17 DE AGOSTO, 12° BPM -BATALHÃO ARRAIAL NOVO DO BOM JESUS, 13° BPM -BATALHÃO CORONEL JOÃO NUNES, 16° BPM -BATALHÃO FREI CANECA, 19° BPM - BATALHÃO ANDRÉ VIDAL DE NEGREIROS) com as localizações das unidades encontradas através do SIG proposto. Nesse sentido, com essa comparação, é possível verificar o desempenho de ambos os cenários e da mesma forma sugerir uma conclusão sobre o modelo proposto.

Logo, para o desenvolvimento e comparação de ambos modelos, foi definido um cenário de cinco instalações com o raio de três quilômetros de cobertura, que de acordo com esses dados de entrada, geraram as informações do Gráfico 15. Em que, é ilustrado a comparação do desempenho do modelo proposto (SIG) com relação a configuração atual das instalações policiais implementada na cidade do Recife (PM).

No Gráfico 15 item A, é apresentado o ganho de cobertura das ocorrências por instalações, ou seja, ao adicionar novas unidades policiais na cidade do Recife, mais crimes são cobertos por essas instalações. E claramente é perceptível o ganho de cobertura a mais que o modelo proposto consegue atingir com relação a configuração atual. No qual, com o modelo proposto é possível atingir ao final da quinta instalação alocada, uma porcentagem de 90% de cobertura das ocorrências, representando uma melhoria de 22% com relação ao cenário atual.

Gráfico 15 – Comparação de desempenho, com relação a cobertura

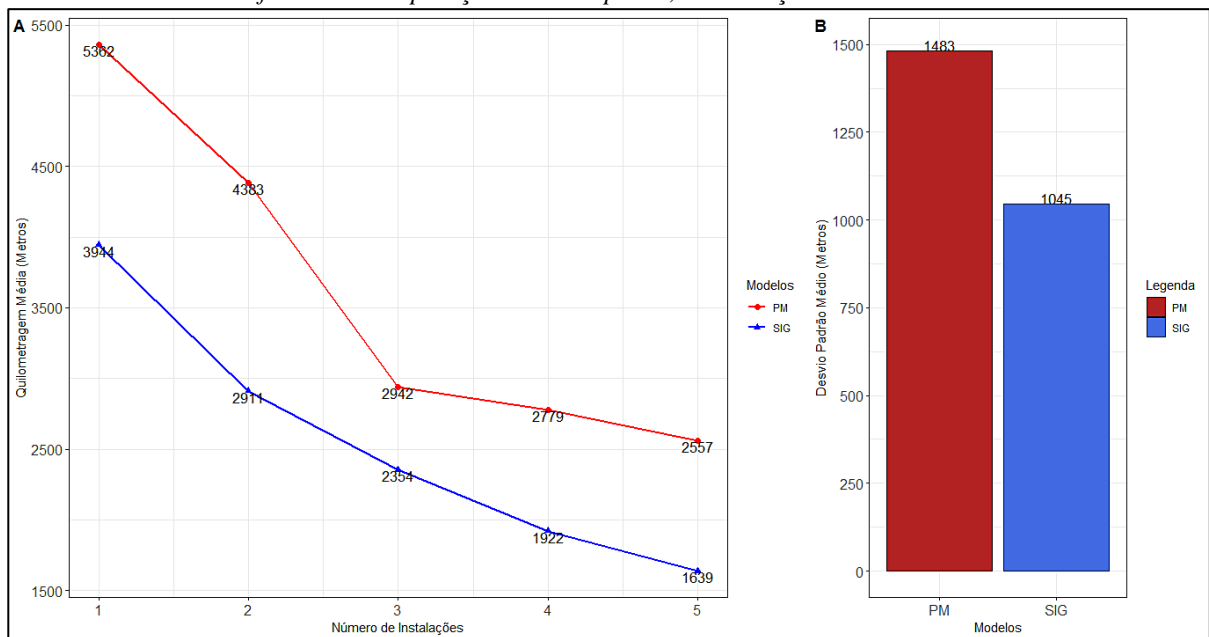


Fonte: O Autor (2019)

Essa melhoria de 22% se torna mais perceptiva ainda, ao analisar esse cenário com relação a quantidade de crimes que são cobertos por essa diferença. Logo, no Gráfico 15 item B, é possível observar a quantidade total de crimes (1688) e as respectivas quantidades em que foram cobertas e não cobertas por cada modelo. Onde com o modelo proposto apenas 167 crimes ficaram descobertos, enquanto que com a configuração atualmente implementada, são 532 crimes descobertos. Diferença essa, bastante considerável, dado a magnitude do problema.

Seguindo com as análises comparativas do mesmo problema, agora por uma perspectiva de distância das instalações até as ocorrências, é ilustrado o Gráfico 16. No qual, é demonstrado a distância média em metros das ocorrências até a instalação mais próxima, de forma que a cada instalação que é adicionada no território do Recife, é obtido uma redução significativa dessa distância.

Gráfico 16 – Comparação de desempenho, com relação a distância média



Fonte: O Autor (2019)

Deste modo no Gráfico 16 item A, é observado que o modelo proposto consegue reduzir de forma significativa a distância média das ocorrências com relação a configuração atual da PM. Conseguindo alcançar uma diferença na quinta instalação alocada de quase 920 metros de distância, o que impacta positivamente no tempo em que as viaturas levam para sair das instalações e chegar nos locais das ocorrências. E para complementar ainda mais essa informação, no Gráfico 16 item B é exibido o desvio padrão médio dessa distância para cada uma das configurações. Mostrando assim, que o modelo proposto consegue estar à frente da configuração atual da PM em todos os cenários aqui avaliados.

Eventualmente, torna-se notório o desempenho do SIG proposto para localizar as instalações policiais na cidade do Recife. Permitindo um melhor desempenho ao cobrir mais ocorrências e reduzir o tempo necessário para que as viaturas alcancem essas ocorrências.

7 CONSIDERAÇÕES FINAIS

Neste capítulo serão apresentadas as considerações finais relacionadas a esta pesquisa. Onde as mesmas, serão apresentadas através da divisão de contribuições da pesquisa, dificuldade e limitações, e as sugestões para trabalhos futuros.

7.1 Contribuições da pesquisa

Como foi analisado na revisão da literatura no capítulo três desta pesquisa, o modelo MCLP tem sua origem em 1974 e atualmente é muito utilizado. Se mostrando um modelo que resistiu a prova do tempo e de fundamental importância na busca da eficiência e otimização dos recursos. Contudo no modelo MCLP, constatou-se a necessidade de um *start point* com relação aos locais candidatos, onde os mesmos são responsáveis por melhorar a eficiência do modelo.

Nesse sentido, como o objetivo desse trabalho foi de criar um SIG que apresentasse um arranjo de localizações que poderiam ser utilizadas para instalar unidades policiais, visando aumentar a eficiência operacional no atendimento às ocorrências. O modelo *K-means* como *start point* na busca por encontrar os locais candidatos, se mostrou um ótimo complemento para o modelo MCLP. E juntos, foram utilizados para o desenvolvimento do SIG proposto nesta pesquisa através do *software RStudio*.

Após o desenvolvimento do SIG proposto e sua aplicação na cidade do Recife- PE, com os dados disponibilizados na plataforma “Onde Fui Roubado”, pode-se dizer que o objetivo geral desta pesquisa foi atingido com sucesso. De modo geral, a pesquisa desenvolveu uma ferramenta inovadora para a modelagem e resolução do problema de localização de instalações policiais.

O modelo proposto no presente estudo obteve desempenho melhor quando comparado com o desempenho da configuração atual das instalações policiais da cidade de Recife. Os resultados obtidos mostraram que o SIG proposto supera a configuração atual das unidades policiais, em termos de localização, em 22% de área coberta e proporcionou uma redução de quase 920 metros na distância média a ser percorrida para atendimento das ocorrências. Em termos práticos, isso significa que as instalações localizadas utilizando a recomendação do SIG proposto, cobrem um maior número de ocorrências e as viaturas percorrem um menor caminho até chegar ao local do crime.

Ademais, vale ressaltar as contribuições sociais do trabalho aqui desenvolvido, ou seja, a relevância dessa pesquisa para a sociedade. Nesse aspecto, reconhece-se que com esta pesquisa,

e seus desenvolvimentos futuros, que é possível melhorar os indicadores de criminalidade a partir da utilização de modelos formais de localização. Assim, é possível obter grandes avanços para ambos os lados, o dos prestadores do serviço em segurança pública e o da população que demanda desses serviços. O presente procedimento de localização possibilita reduzir a quantidade de instalações a ser alocada e maximizar a população que vai ser atendida com um melhor nível de serviço ao minimizar o tempo de espera por uma viatura.

7.2 Dificuldades e limitações

Apesar dos bons resultados obtidos através do desenvolvimento e aplicação do SIG proposto, é importante frisar as dificuldades enfrentadas no seu desenvolvimento e suas limitações atuais.

Nesse sentido, é possível destacar como dificuldade inicial e principal, a carga que é exigida para sua programação computacional. No qual, exige de um certo nível de conhecimento em lógica de programação e a compreensão de alguns algoritmos complexos. E é apenas com esses requisitos mínimos, que se torna possível prosseguir para interpretar, complementar e utilizar as bibliotecas exigidas para sua construção, dentro do ambiente virtual do *software RStudio*.

Já com relação as limitações atuais do SIG, é possível destacar a forma como o mesmo está sendo executado, ainda necessitando de ser executado em conjunto com o *software RStudio*. O que pode ser melhorado, através de plataformas virtuais compatíveis com o próprio *RStudio*. Onde essas plataformas possibilitam alocar o SIG em um sistema de nuvem para ser acessado através de uma página HTML livremente por qualquer pessoa, por meio de computadores e *smartphones*.

Logo, muitas melhorias ainda podem ser agregadas na funcionabilidade e acessibilidade do SIG proposto, principalmente com relação ao seu nível hierárquico no qual é utilizado para auxiliar na tomada de decisão. O que exige, um trabalho contínuo de seu desenvolvimento para garantir o seu aperfeiçoamento no momento de apresentar uma sugestão de localização para as unidades policiais.

7.3 Trabalhos futuros

Dessa forma, a metodologia adotada nesta pesquisa, bem como o sistema de informação gerencial que foi desenvolvido, mostrou-se adequados para de auxiliar na localização de unidades policiais. Principalmente pelos resultados que foram encontrados e comprovados

através da pesquisa observacional e as análises de sensibilidade, o que valida desta forma o estudo.

Sugere-se, entretanto, a continuação desta pesquisa através de uma nova linguagem de programação, mais especificamente a linguagem Python. No qual, é uma linguagem de programação de alto nível, que possui um modelo de desenvolvimento comunitário e aberto, permitindo explorar mais recursos e possivelmente comercializar a ferramenta.

Nesse sentido, os próximos passos devem procurar aplicar o sistema de informação gerencial em outros locais com novas problemáticas de localização. E uma dessas problemáticas que ainda pode ser melhor explorada na cidade do Recife é a consideração de uma sexta localização e compará-la com o desempenho da localização da Central de Plantões da Capital (CEPLANC). Pelo fato de a CEPLANC procurar centralizar os flagrantes realizados pela Polícia Militar do Recife, visando desafogar as delegacias de plantão, que continuam a funcionar normalmente, só que voltadas para as ocorrências da população.

Assim, novas necessidades podem surgir e possibilitar uma melhoria contínua do SIG proposto. Logo, novos módulos podem ser adicionados ao sistema com objetivo de ampliar suas funções, permitindo que outras informações sejam apresentadas, agilizando cada vez mais o trabalho dos tomadores de decisão, no momento de alocar novas unidades policiais.

Os novos módulos do SIG podem se concentrar inicialmente em expandir as possibilidades para encontrar os locais candidatos. Ou seja, incluir na interface uma opção de escolha para que o usuário defina qual método vai ser utilizado para encontrar seus locais candidatos. Esses módulos podem incluir além do modelo *kmeans*, modelos de *p-median problems*, opiniões de especialistas, locais com viabilidade financeira, etc.

Como um passo mais além, seria possível adicionar nas configurações do SIG, modificações que permitissem que o SIG não se limitasse apenas em localizar instalações policiais. Permitindo ao usuário escolher que tipo de variável o mesmo desejasse alocar, sejam elas, escolas, hospitais, pontos de coleta seletiva, base de helicóptero, bases de serviços médicos e equipamentos médicos, bases de ambulâncias, base de bombeiros, base de resgate, instalações postais, antenas TV, antenas Celular, antenas RFID, redes de banda larga, reservas ambientais, torres de vigilância, sirenes de emergência, sensores de segurança, anúncios em outdoor, paradas de ônibus e etc.

Por fim, como incremento final para o SIG seria possível anexar uma página de análise financeira. Onde essa página seria responsável por realizar estudos de viabilidade sobre os locais encontrados. Pois cada local encontrado pelo SIG pode exigir de um valor monetário diferente, o que vai influenciar na decisão final do usuário.

Nesse sentido, com a adição de todos esses novos módulos e outros mais que fossem surgindo através da elaboração progressiva do SIG, seria possível incrementar e envolver mais as preferências do decisor dentro do modelo. Evoluindo assim, o sistema de informação gerencial para um sistema de apoio a decisão. No qual, pode oferecer mais recursos e ser utilizado em uma nova posição hierárquica dentro de uma organização.

REFERÊNCIAS

- AQUINO A. C. B.; PAGLIARUSSI M. S.; BITTI E. J. S. Heurística para a composição de referencial teórico. *Revista Contabilidade & Finanças*, São Paulo, v. 19, n. 47, 2008.
- A. T. Murray. Maximal coverage location problem: impacts, significance, and evolution. *International Regional Science Review*, 39(1):5–27, 2016.
- A. WEBER, Uber den Standort der Industrien. *Alfred Weber's Theory of the Location of Industries*, University of Chicago, 1929.
- AGUADO, J. S.; TRANDAFIR, P. C. Some heuristic methods for solving p-median problems with a coverage constraint, *European Journal of Operational Research*, Volume 220, Issue 2, Pages 320-327, 16 July, 2012.
- ALCARAZ, J.; LANDETE, M.; MONGE, J. F. Design and analysis of hybrid metaheuristics for the Reliability p-Median Problem. *European Journal of Operational Research*, v. 222, n. 1, p. 54–64, 2012.
- ALUMUR, S.; KARA, B. Y. Network hub location problems: The state of the art. *European Journal of Operational Research*, v. 190, n. 1, p. 1–21, 2008.
- ALY, A. A.; LITWHILER, D. W. *Police Briefing Stations: A Location Problem*. AIIE Transactions, v. 11, n. 1, p. 12–22, 1979.
- ANDRADE E. L. *Introdução à pesquisa operacional: métodos e modelos para análise de decisões*. – 4.ed. Rio de Janeiro: LTC, 2009.
- AN, H. C. et al. Centrality of trees for capacitated k-center. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 8494 LNCS, p. 52–63, 2014.
- ANGELONI, M. T. *Elementos intervenientes na tomada de decisão*. Ci. Inf., Brasília, v. 32, n. 1998, p. 17–22, 2003.
- ARIA, M. e CUCCURULLO, C. *Bibliometrix: an R-tool for comprehensive science mapping analysis*. J. Informetr 11. 959–975, 2017.

BANDYOPADHYAY, M.; SINGH, M.; SINGH, V. Spatial Pattern Analysis for finding Weighted Candidate Set for p-median Problem in Locating Emergency Facilities. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, Volume 2, Issue 5, May 2012.

BALLOU, R. H. *GERENCIAMENTO DA CADEIA DE SUPRIMENTOS/LOGÍSTICA EMPRESARIAL*. 5. ed., Porto Alegre, Bookman, 2006.

BANSAL, M.; KIANFAR, K. Planar maximum coverage location problem with partial coverage and rectangular demand and service zones. *INFORMS Journal on Computing*, v. 29, n. 1, p. 152–169, 2017.

BAO, S. *et al.* Optimizing watchtower locations for forest fire monitoring using location models. *Fire Safety Journal*, v. 71, n. December 2013, p. 100–109, 2015.

BERMAN, O.; HUANG, R. The minimum weighted covering location problem with distance constraints. *Computers and Operations Research*, v. 35, n. 2, p. 356–372, 2008.

BERKHIN, P. “*Survey of Clustering Data Mining Techniques*”, Accrue Software, Inc. 2002.

BRERETON, P.; KITCHENHAM, B.; BUDGEN D.; TURNER, M.; KHALIL, M. “Lessons from applying the systematic literature review process within the Software Engineering domain”. *The Journal of Systems and Software*, 80, pp. 571–583. 2007.

BROWN K, COULTER P. B. *Subjective and objective measures of police service delivery*. American Society for Public Administration. v. 43, n. 1, p. 50–58, 1983.

BORGES, G. R.; FERNANDES, F. C. A Contribuição Do Sistema De Informação Gerencial Para A Obtenção De Vantagem Competitiva. *Revista Sinergia*, Rio Grande, 20 (1): 29-40, 2016.

CARDOSO, R. S. *et al.* USO DE SAD NO APOIO À DECISÃO NA DESTINAÇÃO DE RESÍDUOS PLÁSTICOS E GESTÃO DE MATERIAIS. *Pesquisa Operacional*, v. 29, p. 67–95, 2009.

CANO, I. “Políticas de segurança pública no Brasil: tentativas de modernização e democratização versus a guerra contra o crime”. *SUR – Revista Internacional de Direitos Humanos*, nº 5, ano 3, p. 136-155, 2006.

CANIËLS, Marjolein CJ; BAKENS, Ralph JJM. The effects of Project Management Information Systems on decision making in a multi project environment. *International journal of project management*, v. 30, n. 2, p. 162-175, 2012.

CAMPEÃO *et al.* *Sistema De Informação Gerencial: Um Modelo Conceitual Para Sistemas Locais De Produção*. XXVII ENEGEP, Foz do Iguaçu, 2007.

CARVALHO, V. A.; SILVA, M. R. F. Política de segurança pública no Brasil: avanços, limites e desafios. *Rev. Katálysis*. Florianópolis, v. 14, n. 1, p. 59-67, junho de 2011.

CLARO, M. A. P. M. e NICKEL, D. C. *Gestão do capital humano*. Fae Business School. Curitiba: Associação Franciscana de Ensino Senhor Bom Jesus, 2002

CERQUEIRA, D. R. C.; CARVALHO, A.; LOBAO, W. J. A.; RODRIGUES, Rute I. *Análise dos custos e consequências da violência no Brasil*. Texto para Discussão n. 1.284. Brasília: IPEA, 2007.

CHIYOSHI, F.; GALVAO, R. D.; MORABITO, R. O uso do modelo hipercubo na solução de problemas de localização probabilística. *Gest. Prod.* São Carlos, v. 7, n. 2, p. 146-174, Agosto de 2000.

CHURCH RL, STOMS DM, DAVIS FW; Reserve selection as a maximal covering location problem. *Biol Conserv* 76:105–112, 1996.

CHURCH, R. L.; REVELLE, C. *The maximal covering location problem*. Papers of the Regional Science Association, 32, p. 101–118, 1974.

CIRINO, S. e *et al.* *Modelo não linear de localização de instalações de serviços de saúde com indicador de acessibilidade incorporado*. Cad. Saúde Pública, Rio de Janeiro, v. 34, n. 3, 2018.

CROMLEYA, R. G *et al.* Evaluating representation and scale error in the maximal covering location problem using GIS and intelligent areal interpolation. *International Journal of Geographical Information Science*, Vol. 26, No. 3, 495–517, 2012.

COHEN-ADDAD, V. *et al.* Local Search Yields Approximation Schemes for k-Means and k-Median in Euclidean and Minor-Free Metrics. *SIAM J. COMPUT.*, v. 48, n. 2, p. 644-667, 2019.

CURRENT JR, OKELLY M, Locating emergency warning sirens. *Decis Sci* 23:221–234,1992

CURTIN, K. M.; HAYSLETT-MCCALL, K.; QIU, F. *Determining Optimal Police Patrol Areas with Maximal Covering and Backup Covering Location Models*. Network and Spatial Economics, 2007.

DANE, F. Research methods. *Brooks/Cole Publishing Company: California*, 1990.

DALFOVO, O.; DALFOVO, L. L. D. Sistema de Informação Gerencial Observatório da Educação. Rev. GUAL., Florianópolis, v.2, n. 1, p.62-77, ano 2009.

DASKIN, M.S., *Network and Discrete Location: Models, Algorithms and Applications*. John Wiley, New York, 2001.

DAVOODI, M. k-Balanced Center Location problem: A new multi-objective facility location problem. *Computers and Operations Research*, v. 105, p. 68–84, 2019.

DAVID, S.; SHWARTZ, S. S. *Understanding Machine Learning: From Theory to Algorithms*, p. 1-449, v.1, 2014.

DHALMAHAPATRA, K. *et al.* Decision support system for safety improvement: An approach using multiple correspondence analysis, t-SNE algorithm and K-means clustering. *Computers and Industrial Engineering*, v. 128, p. 277–289, 2019.

DIAS NETO, Theodomiro. *Segurança urbana: o modelo da nova prevenção*. São Paulo: Editora Revista dos Tribunais: Fundação Getúlio Vargas, 2005.

EISELT, H.A.; LAPORTE, G. Objectives in location problems. In: Drezner, Z. (Ed.), *Facility Location: A Survey of Applications and Methods*. Springer-Verlag, Berlin, p. 151–180, 1995.

ERDEMIR, E.T., et al. Joint ground and air emergency medical services coverage models: a greedy heuristic solution approach. *European Journal of Operational Research*, 207, 736–749, 2010.

FALLAH, H.; NAIMISADIGH, A.; ASLANZADEH, M. *Covering problem*. In *Facility location: Concepts, models, algorithms and case studies*. Heidelberg, Germany: Physica Verlag, 2009.

FARAHANI, R. Z. *et al.* Covering problems in facility location: A review. *Computers and Industrial Engineering*, v. 62, n. 1, p. 368–407, 2012.

FARAHANI, R. Z. *et al.* Hierarchical facility location problem: Models, classifications, techniques, and applications. *Computers and Industrial Engineering*, v. 68, n. 1, p. 104–117, 2014.

FARAHANI, R. Z. *et al.* OR models in urban service facility location: A critical review of applications and future developments. *European Journal of Operational Research*, v. 276, n. 1, p. 1–27, 2019.

FARAHANI, R.Z., STEADIESEIFI, M., ASGARI, N., Multiple criteria facility location problems: a survey. *Appl. Math. Modell.* 34 (7), 1689–1709, 2010.

FAZEL ZARANDI, M. H.; DAVARI, S.; HADDAD SISAKHT, S. A. The large scale maximal covering location problem. *Scientia Iranica*, v. 18, n. 6, p. 1564–1570, 2011.

FERNANDES, S.; CAPTIVO, M. E.; CLIMACO, J. SABILOC – UM SISTEMA DE APOIO À DECISÃO PARA ANÁLISE DE PROBLEMAS DE LOCALIZAÇÃO BICRITÉRIO *Pesqui. Oper.*, Rio de Janeiro, v. 27, n. 3, p. 607-628, 2007.

FERREIRA, D. C. *et al.* Avaliação do Sistema de Informações Gerenciais da Assembleia Legislativa do Amazonas. *Contabilidade, Gestão e Governança*, v. 18, n. 2, 2015.

FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA, Anuário Brasileiro de Segurança Pública. 10º Edição, ISSN 1983-7364, São Paulo: Fórum Brasileiro de Segurança Pública, 2016. Disponível em: <http://www.forumseguranca.org.br/publicacoes/10o-anuario-brasileiro-de-seguranca-publica>

FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA, Anuário Brasileiro de Segurança Pública. 11º Edição, ISSN 1983-7364, São Paulo: Fórum Brasileiro de Segurança Pública, 2017. Disponível em: <http://www.forumseguranca.org.br/publicacoes/11o-anuario-brasileiro-de-seguranca-publica>

FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA, Anuário Brasileiro de Segurança Pública. 12º Edição, ISSN 1983-7364, São Paulo: Fórum Brasileiro de Segurança Pública, 2018. Disponível em: <http://www.forumseguranca.org.br/publicacoes/anuario-brasileiro-de-seguranca-publica-2018>

FRANCIS, R. L.; WHITE, J. A. *Facility layout and location an analytical approach*, 1st ed. Englewood Cliffs, NJ, US: Prentice-Hall, 1974.

FRIGGSTAD, Z. *et al.* *Approximation Schemes for Clustering with Outliers*. v. 15, n. 2, 2017.

FRIGGSTAD, Z. *et al.* LOCAL SEARCH YIELDS A PTAS FOR k-MEANS IN DOUBLING METRICS. *SIAM J. COMPUT*, v. 48, n. 2, p. 452–480, 2019.

GALDINO, J. A.; GUIMARÃES, M. DA G. V.; FILHO, M. M. DO C. Análise Do Desempenho Na Gestão Das Despesas Orçamentárias Com Segurança Pública No Brasil. *Revista Ambiente Contábil*, v. 9, n. 92, p. 7–9, 2018.

GARLAND, David. *A cultura do controle: crime e ordem na sociedade contemporânea*. Rio de Janeiro: Revan, 2014.

GARNER, A. A.; VAN DEN BERG, P. L. Locating helicopter emergency medical service bases to optimise population coverage versus average response time. *BMC Emergency Medicine*, v. 17, n. 1, p. 1–11, 2017.

GELAREH, S.; NICKEL, S. *Hub location problems in transportation networks*. *Transportation Research Part E: Logistics and Transportation Review*, v. 47, n. 6, p. 1092–1111, 2011.

GIL, Antonio Carlos. *Como elaborar projetos de pesquisa*. 5ª ed. São Paulo, SP.: Atlas, 2010

GRUBESIC, T. H.; MURRAY, A. T. Constructing the divide: Spatial disparities in broadband access. *Papers in Regional Science*, v. 81, n. 2, p. 197–221, 2002.

GUAZZELLI, C. S.; CUNHA, C. B. Multi-criteria optimization for the location problem of distribution centers of a company with production plant in the Industrial Pole of Manaus. *Gest. Prod.*, São Carlos, v. 22, n. 3, p. 480-494, Sept. 2015.

GUNAWARDANE, G. Dynamic versions of set covering type public facility location problems. *European Journal of Operational Research*, v. 10, n. 2, p. 190–195, 1982.

GUSMÃO, A. P. H. *et al.* *The Use of a Decision Support System to Aid a Location Problem Regarding a Public Security Facility*. P. S. A. Freitas et al. (Eds.): EmC-ICDSST 2019, LNBIP 348, pp. 15–27, 2019.

HAIR, J. F. ROLPH, A. E. TATHAM, R. L. BLACK, W. C. *Análise Multivariada de Dados*, 5 ed, Bookman, Porto Alegre, Brasil, 2005

HAKIMI, S.L., Optimum locations of switching centers and the absolute centers and medians of a graph, *Operations Research* 12, p. 450–459., 1964.

HARTIGAN, J. A. *Clustering Algorithms*, New York: John Wiley and Sons, 113–129, 1975.

HU DING; LUNJIA HU; LINGXIAO HUANG; JIAN LI; *Capacitated Center Problems with Two-Sided Bounds and Outliers*. Springer International Publishing AG 2017, F. Ellen et al. (Eds.): WADS 2017, LNCS 10389, pp. 325–336, 2017.

HUANG, M. et al. Interest-Driven Outdoor Advertising Display Location Selection Using Mobile Phone Data. *IEEE Access*, v. 7, p. 30878–30889, 2019.

IPEA. Atlas da Violência, *Instituto de Pesquisa Econômica Aplicada*, 2018

J. REESE, *Solution methods for the p-median problem: an annotated bibliography*, Networks 48 125–142, 2006.

JEAN-FRANC, OIS CORDEAU, FABIO FURINI, IVANA LJUBIC, Benders Decomposition for Very Large-Scale Partial Set Covering and Maximal Covering Location Problems, *European Journal of Operational Research*, 2018.

JUNG W., *A review of research: an investigation of the impact of data quality on decision performance*. In: International Symposium on Information & Communication Technologies (ISITC'04), 2004.

KANUNGO, T. et al. *An efficient k-means clustering algorithm: Analysis and implementation*, IEEE transactions on pattern analysis and machine intelligence, 881–892, 2002.

KAVEH, P. ZADEH, S. A. SAHRAEIAN, R. Solving Capacitated P-median Problem by Hybrid K-means Clustering and FNS Algorithm. *International Journal of Innovation, Management and Technology*. Industrial Engineering and Operations Management, Dhaka, Bangladesh, v. 1, n. 4, 2010.

KCHAOU BOUJELBEN, M.; BOULAKSIL, Y. *Modeling international facility location under uncertainty: A review, analysis, and insights*. IISE Transactions, v. 50, n. 6, p. 535–551, 2018.

KRIVULIN, N. Using tropical optimization to solve constrained minimax single-facility location problems with rectilinear distance. *Computational Management Science*, v. 14, n. 4, p. 493–518, 2017.

KOO, C. *et al.* A novel estimation approach for the solar radiation potential with its complex spatial pattern via machine-learning techniques. *Renewable Energy*, p. 575–592, 2019.

LANTZ, B. *Machine Learning with R*. Discover how to build machine learning algorithms, prepare data, and dig deep into data prediction techniques with R, ed. 2, p 1- 452, 2015.

LAUDON, Kenneth C.; LAUDON, Jane Price. *Sistemas de informação gerenciais: administrando a empresa digital*. São Paulo: Prentice Hall, 2004.

LI, Chan et al. The consequences of information technology control weaknesses on management information systems: The case of Sarbanes-Oxley internal control reports. *MIS Quarterly*, v. 36, n. 1, p. 179-204, 2012.

LIMA, Renato Sérgio de; BUENO, Samira; MINGARDI, Guaracy. Estado, polícias e segurança pública no Brasil. *Rev. direito GV*, São Paulo, v. 12, n. 1, p. 49-85, Apr. 2016.

LIMA, Renato Sérgio de; SINHORETTO, Jacqueline; BUENO, Samira. *A gestão da vida e da segurança pública no Brasil*. Soc. estado., Brasília, v. 30, n. 1, p. 123-144, Apr. 2015.

LIU, M. *et al.*, *Online data organizer: micro-video categorization by structure-guided multimodal dictionary learning*, IEEE Trans. Image Process, 1235–1247, 2019.

LORAAMM, R. W.; DOWNS, J. A. A wildlife movement approach to optimally locate wildlife crossing structures. *International Journal of Geographical Information Science*, v. 30, n. 1, p. 74–88, 2016.

LORENA, L. A.N, *Análise Espacial de Redes com Aplicações em Sistemas de informações Geográficas*, INPE, São José dos Campos, 2003.

MACQUEEN, J. “*Some Methods for Classification and Analysis of Multivariate Observations.*” In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1(14), 281–297, 1967.

MARCONI, M. A.; LAKATOS, E. M. *Pesquisa bibliográfica e resumos*. In: fundamentos de metodologia científica. 7ª edição – São Paulo: Atlas, 2010.

MAPA, Sílvia Maria Santana; LIMA, Renato da Silva. Combining geographic information systems for transportation and mixed integer linear programming in location-allocation problems. *Gest. Prod.* São Carlos, v. 19, n. 1, p. 119-136, 2012.

MAKINDE *et al.* *Lessons learned in the deployment of a HIV counseling and testing management information system on a new project*. *AIDS*, v. 25, p. 2289–2293, 2011.

MÁXIMO, V. R.; NASCIMENTO, M. C. V.; CARVALHO, A. C. P. L. F. Intelligent-guided adaptive search for the maximum covering location problem. *Computers and Operations Research*, v. 78, p. 129–137, 2017.

MITCHELL P. Optimal selection of police patrol beats. *The Journal of Criminal Law, Criminology and Police Science*. Vol. 63, n. 4. 1972.

MITHAS, Sunil; RAMASUBBU, Narayan; SAMBAMURTHY, Vallabh. How information management capability influences firm performance. *MIS quarterly*, v. 35, n. 1, p. 237-256, 2011.

MORAIS GURGEL, A.; PIRES FERREIRA, R.J.; ALOISE, D.J. *Proposta de modelos para a localização de bases policiais e roteirização de viaturas*. ENEGEP. 2010.

MONACO, Kristen; REDMON, Brydey. Does contracting with owner operators lead to worse safety outcomes for US motor carriers? Evidence from the Motor Carrier Management Information System. *Accident Analysis & Prevention*, v. 45, p. 654-659, 2012.

MURRAY, A. T. *et al.* Coverage optimization to support security monitoring. *Computers, Environment and Urban Systems*, v. 31, n. 2, p. 133–147, 2007.

NEEMA, M. N.; MANIRUZZAMAN, K. M.; OHGAI, A. New Genetic Algorithms Based Approaches to Continuous p-Median Problem. *Networks and Spatial Economics*, v. 11, n. 1, p. 83–99, 2011.

NÓBREGA JR, J. M. N.; ZAVERUCHA, J. Violência homicida no nordeste brasileiro: uma refutação às explicações baseadas na desigualdade e na pobreza. *Anuário Antropológico*, n. II, p. 53–87, 2017.

NÓBREGA JR., José Maria Pereira da. Violência homicida no Nordeste brasileiro: Dinâmica dos números e possibilidades causais. *DILEMAS: Revista de Estudos de Conflito e Controle Social* – Vol. 10 – no 3, pp. 553-572, 2017.

NÓBREGA JÚNIOR, J. M. P. DA. Diagnóstico da violência no Brasil e os Desafios para a Segurança Pública. *Revista Espaço Acadêmico*, v.14, n. 167, p. 103–115, 2015.

OLIVEIRA, D. P. R. *Sistemas de informações gerenciais: estratégias, táticas, operacionais*. ed. 9, São Paulo: Atlas, 2004.

OLIVEIRA, M G. *Sistema de Localização de Facilidades: Uma abordagem para mensuração de pontos de demanda e localização de facilidades*, Dissertação apresentada ao Programa de Pós-Graduação do Instituto de Informática da Universidade Federal de Goiás, Goiânia, 2012.

OLIVEIRA, L. M. B. e MORAES, W. F. A. Coleta de dados realizada por questionário enviado pelo correio: método eficaz? *Rev. adm. empres.*, São Paulo, v. 34, n. 4, p. 85-92, 1994.

O'KELLY, M.E., A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research* 32, 393–404, 1987.

OTTO, B.; BOYSEN, N. A dynamic programming-based heuristic for locating stops in public transportation networks. *Computers and Industrial Engineering*, v. 78, p. 163–174, 2014.

OYELADE, O. J.; OLADIPUPO, O. O.; OBAGBUWA, I. C. Application of k Means Clustering algorithm for prediction of Students Academic Performance. *International Journal of Computer Science and Information Security*, v. 7, p. 292–295, 2010.

OWEN, S. H.; DASKIN, M. S. Strategic facility location: a review. *European Journal of Operational Research*, n. 111, p. 423-447, 1998.

OZTEKIN, A. et al. Criticality index analysis based optimal RFID reader placement models for asset tracking. *International Journal of Production Research*, v. 48, n. 9, p. 2679–2698, 2010.

PLOSKAS, N. et al. *An Interactive Spatial Decision Support System Enabling Co-Located Collaboration using Tangible User Interfaces for the Multiple Capacitated Facility Location Problem*. *International Journal of Decision Support System Technology*, v. 7, n. 2, p. 15–28, 2015.

- PRAVEEN KUMAR, D.; AMGOTH, T.; ANNAVARAPU, C. S. R. *Machine learning algorithms for wireless sensor networks: A survey. Information Fusion*, v. 49, p. 1–25, 2019.
- PORTO, M. A. G.; BANDEIRA, A. A. *A importância dos sistemas de informações gerenciais para as organizações. XIII SIMPEP - Bauru, SP, Brasil, 6 a 8 de novembro de 2006.*
- PIZZOLATO, N. D. RAUPP, F. M. P. Alzamora, G. S. *Revisão de desafios aplicados em localização com base em modelos da p-mediana e suas variantes, SOBRAPO, Rio de Janeiro, Brasil. v.4, n.1, p. 13-42., 2012.*
- PULVER, A.; WEI, R.; MANN, C. Locating AED enabled medical drones to enhance cardiac arrest response times. *Prehospital Emergency Care*, v. 20, n. 3, p. 378–389, 2016.
- REVELLE C, SWAIN R. *Central Facilities Location. Geogr Anal* 2:30–42, 1970.
- REVELLE, C. S. EISELT, H.A. Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165. 1–19, 2005.
- REVELLE, C. S.; MARKS, D.; LIEBMAN, J. Analysis of private and public sector location problems. *Management Science* 16, p. 692–707, 1970.
- RISSO, L. A. et al. *Clusterização K-Means Para Decisão De Localização Industrial. XXXV ENCONTRO NACIONAL DE ENGENHARIA DE PRODUCAO, Fortaleza, CE, 2015.*
- RØISLIEN, J. et al. Exploring optimal air ambulance base locations in Norway using advanced mathematical modelling. *Injury Prevention*, v. 23, n. 1, p. 10–15, 2017.
- GARCÍA, S.; MARÍN, A. *Covering location problems. In Location Science*, pages 93–114. Springer, 2015.
- SANTOS, Iris Gomes dos; GONTIJO, José Geraldo Leandro; AMARAL, Ernesto F. L. A política de segurança pública no Brasil: uma análise dos gastos estaduais (1999-2010). *Opin. Publica*, Campinas, v. 21, n. 1, p. 105-131, Apr. 2015.
- SENTO-SÉ, J. T. *Prevenção da violência: o papel das cidades. Rio de Janeiro: Civilização Brasileira, 2006.*
- SILVA, C. C. C. ALTERNATIVAS À POLICIALIZAÇÃO DA SEGURANÇA PÚBLICA. *REVISTA ESMAT*, [S.l.], v. 2, n. 2, p. 47-64, ISSN 2447-9896, abr. 2017.

SILVA, V. F.; CABRAL, L. A. F.; QUIRNO, R. *Estratégia para seleção de locais candidatos voltada ao problema de localização de viaturas policiais*. XLVSBPO- 19 Simpósio Brasileiro de Pesquisa Operacional, Natal/RN, setembro de 2013.

SNYDER, S. A.; HAIGHT, R. G. Application of the Maximal Covering Location Problem to Habitat Reserve Site Selection: A Review. *International Regional Science Review*, v. 39, n. 1, p. 28–47, 2016.

SNYDER, L. V. *Covering problems*. In *Foundations of Location Analysis*, pages 109–135. Springer, 2011.

STAIR, R. M.; REYNOLDS, G. W. *Princípios De Sistemas De Informação*. Editora: Cengage Learning, ISBN: 8522118620, 2015.

SOARES, J. L. *Prevenção da violência: o papel das cidades*. Rio de Janeiro: Civilização Brasileira, 2006.

SOHNGEN, C.; CIPRIANI, M. Crime E Violência No Brasil: Representações Socioculturais Na Pós-Modernidade. *R. Inter. Interdisc. INTERthesis*, Florianópolis, v.16, n.1, p. 36-56 Jan-Abr 2019.

STEFANELLO, F.; DE ARAÚJO, O. C. B.; MÜLLER, F. M. Matheuristics for the capacitated p-median problem. *International Transactions in Operational Research*, v. 22, n. 1, p. 149–167, 2015.

TAYLOR P.E; HUXLEY S.J. *A break from tradition for the San-Francisco police-patrol officer scheduling using an optimization-based decision support system*. *Interfaces* 19:4–24, 1989.

TELLA, R. D.; EDWARDS, S.; SCHARGRODSKY, E. *Introduction*. In: *The economics of crime: lessons for and from Latin America*. Chicago: University of Chicago Press, 2010.

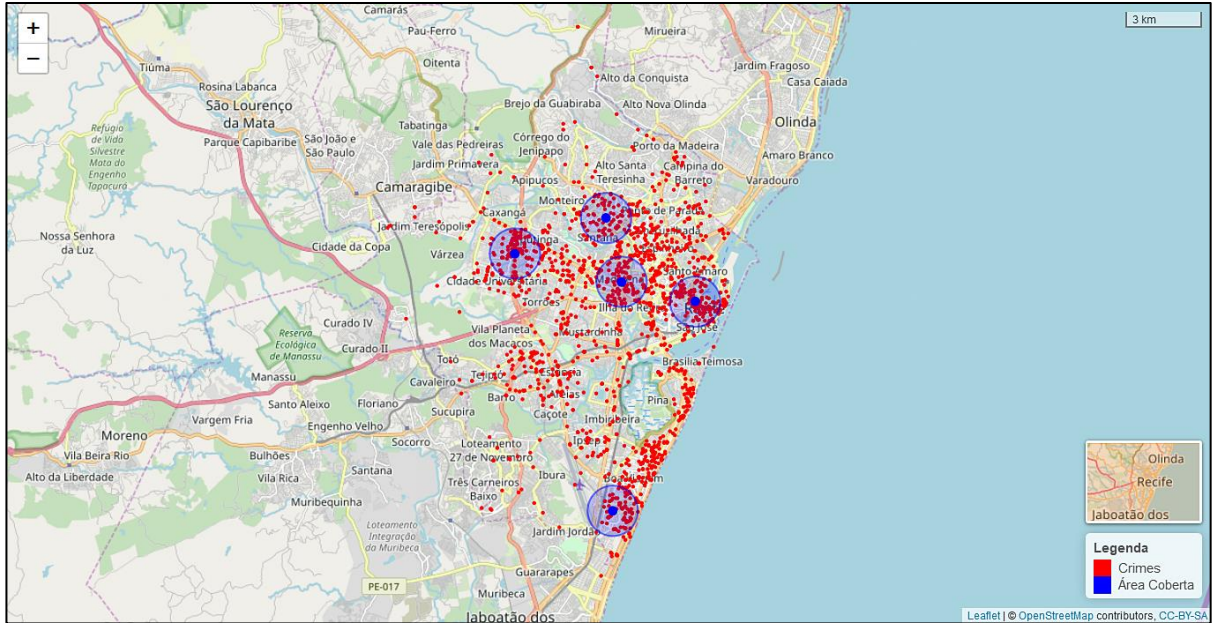
TOREGAS, C. et al. The Location of Emergency Service Facilities. *Operations Research*, v. 19, n. 6, p. 1363–1373, 1971.

TRINDADE, A. Estado, governança e segurança pública no Brasil: Uma análise das secretarias estaduais de Segurança Pública. *Dilemas*, v. 8, p. 607–632, 2015.

- TURBAN, E.; ARONSON, J. E.; LIANG, T.-P. *Decision Support System and Intelligent System*. Prentice Hall, 7th ed., 2005.
- TURRIONI, J. B; MELLO, C. H. P. *Metodologia De Pesquisa Em Engenharia De Produção*, UNIFEI, p.83-85, 2012.
- VATSA, A. K.; JAYASWAL, S. A new formulation and Benders decomposition for the multi-period maximal covering facility location problem with server uncertainty. *European Journal of Operational Research*, v. 251, n. 2, p. 404–418, 2016.
- WAISELFISZ, J. J. *Mapa da violência 2014: os jovens do Brasil*. Brasília, DF: Centro Brasileiro de Estudos Latino-Americano, 2014.
- WANG, J. et al. A new partial coverage locating model for cooperative fire services. *Information Sciences*, v. 373, p. 527–538, 2016.
- WANG, K. et al. Optimizing the configuration of precipitation stations in a space-ground integrated sensor network based on spatial-temporal coverage maximization. *Journal of Hydrology*, v. 548, p. 625–640, 2017.
- WEBSTER, J.; WASTON, J.T. Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly & The Society for Information Management*, v.26, n.2, pp.13-23, 2002
- YIN, P.; MU, L. An empirical comparison of spatial demand representations in maximal coverage modeling. *Environment and Planning B: Planning and Design*, v. 42, n. 4, p. 574–592, 2015.
- YIN, P.; MU, L. Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles. *Applied Geography*, v. 34, p. 247–254, 2012.
- ZAVERUCHA, Jorge; NÓBREGA JR., José Maria Pereira da. (2015), “O pacto pela vida, os tomadores de decisão e a redução da violência homicida em Pernambuco”. *Dilemas: Revista de Estudos de Conflito e Controle Social*, Vol. 8, nº 2, pp. 235-252.
- ZHAO, Y.-L. et al., *Personalized recommendations of locally interesting venues to tourists via cross-region community matching*, ACM Trans. Intell. Syst.Technol. 2014.
- ZHOU, X. et al., Integrating Island Spatial Information and Integer Optimization for Locating Maritime Search and Rescue Bases: A Case Study in the South China Sea. *ISPRS International Journal of Geo-Information*, v. 8, n. 2, p. 88, 2019.

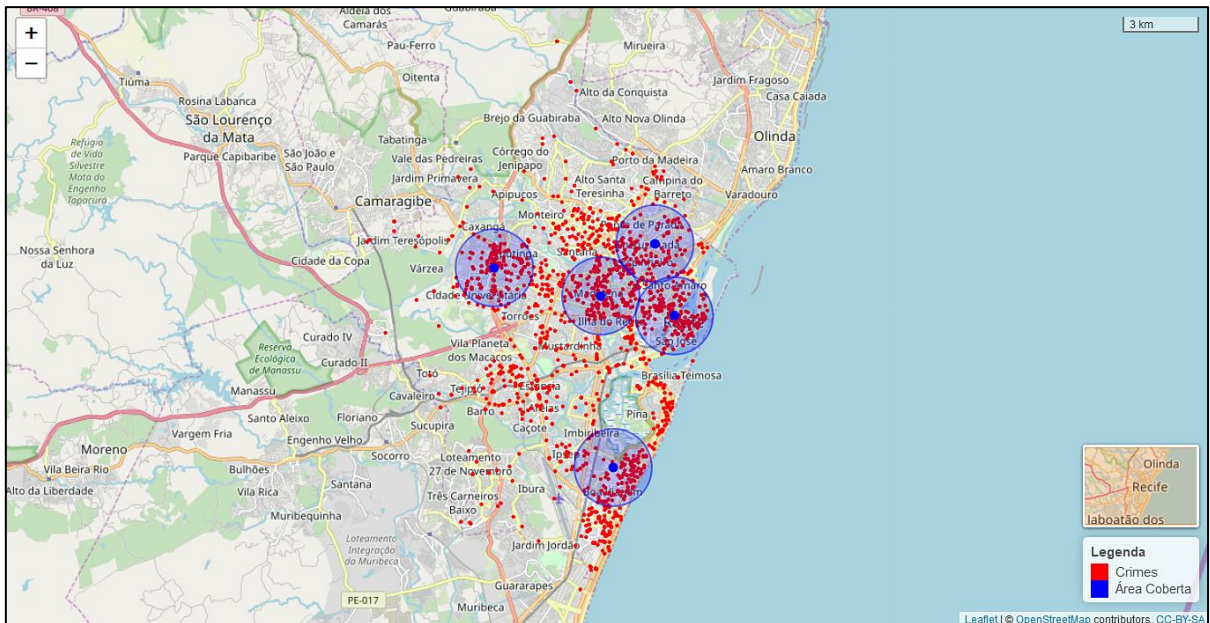
APÊNDICE A – RESULTADOS DA ALOCAÇÃO OTIMIZADA

Mapa 6 – Instalações ótimas alocadas, com raio de 1.0 km de cobertura



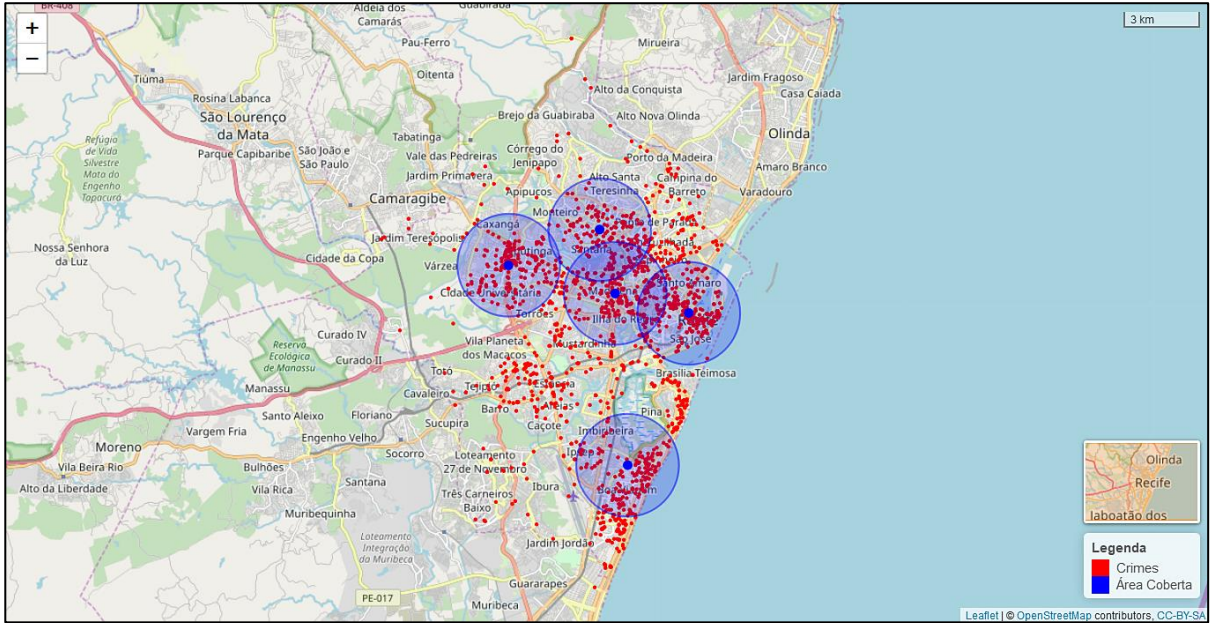
Fonte: O Autor (2019)

Mapa 7 – Instalações ótimas alocadas, com raio de 1.5 km de cobertura



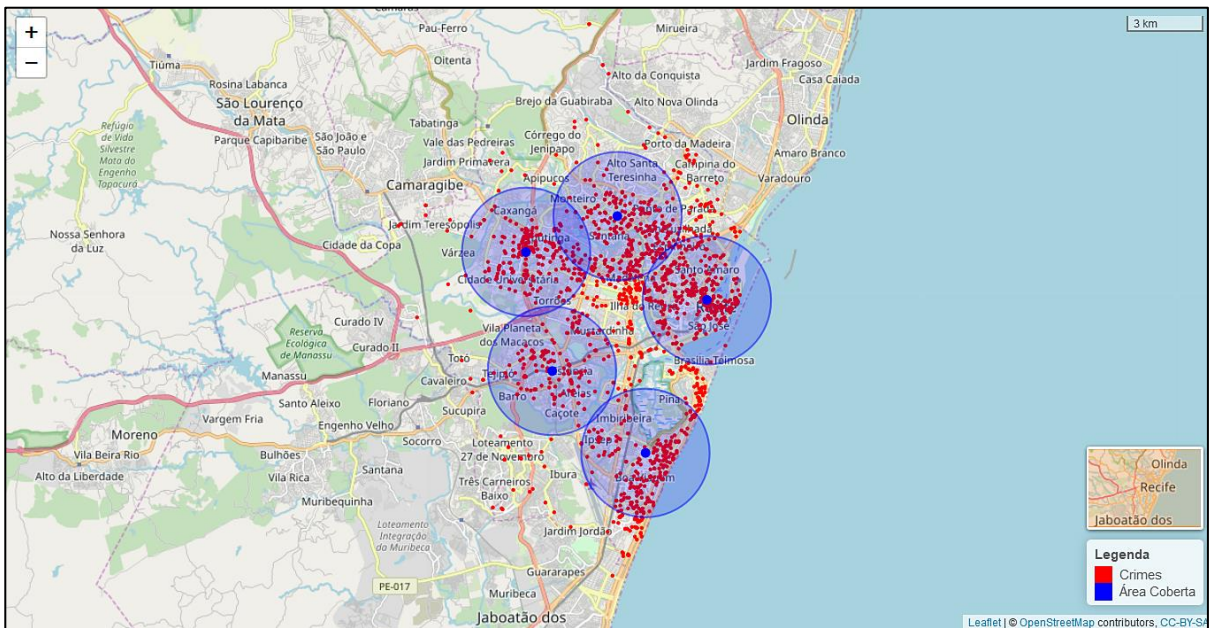
Fonte: O Autor (2019)

Mapa 8 – Instalações ótimas alocadas, com raio de 2.0 km de cobertura



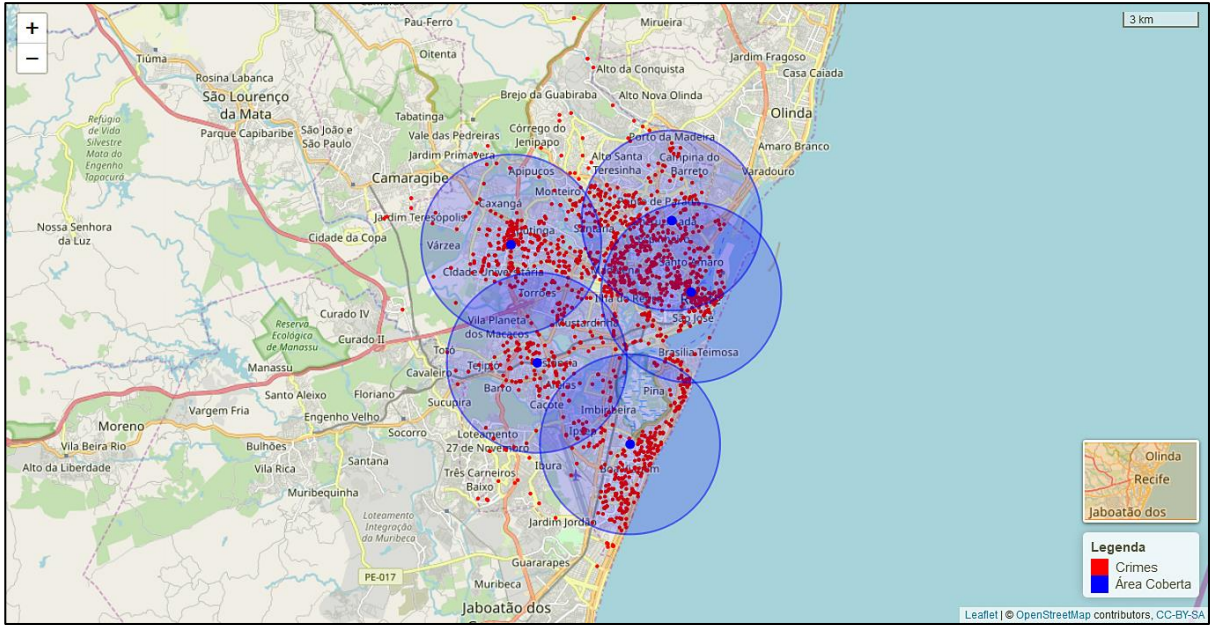
Fonte: O Autor (2019)

Mapa 9 – Instalações ótimas alocadas, com raio de 2.5 km de cobertura



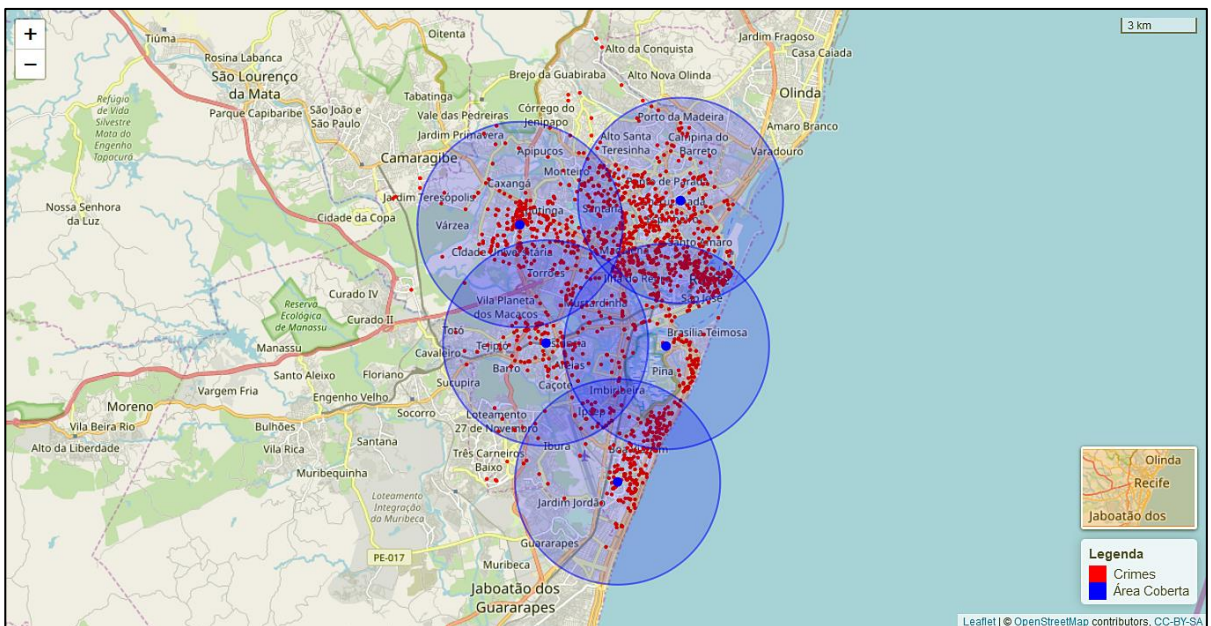
Fonte: O Autor (2019)

Mapa 10 – Instalações ótimas alocadas, com raio de 3.5 km de cobertura



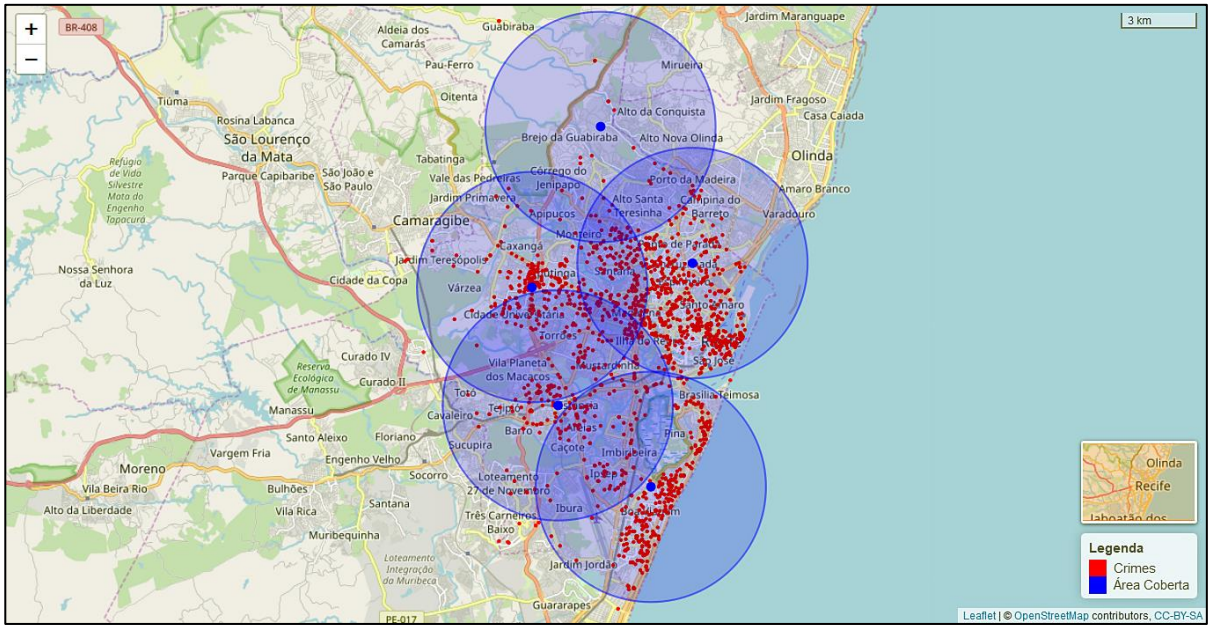
Fonte: O Autor (2019)

Mapa 11 – Instalações ótimas alocadas, com raio de 4.0 km de cobertura



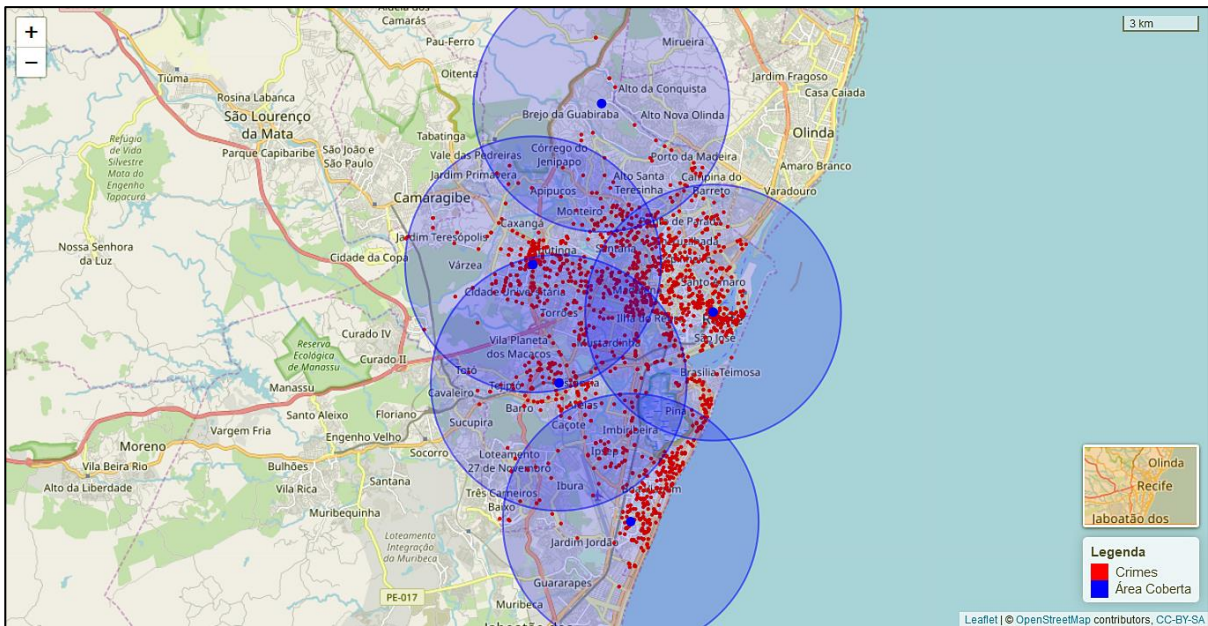
Fonte: O Autor (2019)

Mapa 12 – Instalações ótimas alocadas, com raio de 4.5 km de cobertura



Fonte: O Autor (2019)

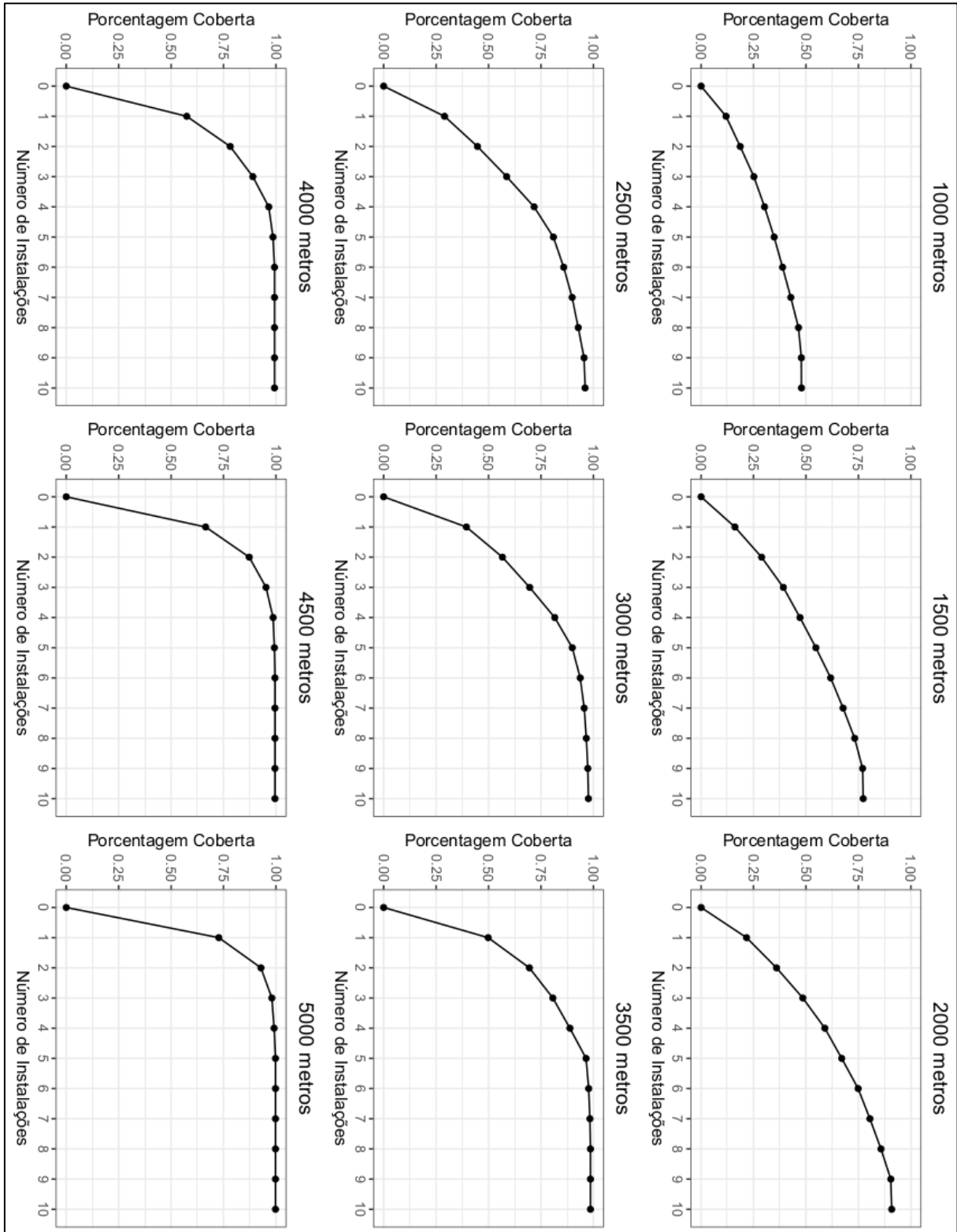
Mapa 13 – Instalações ótimas alocadas, com raio de 5.0 km de cobertura



Fonte: O Autor (2019)

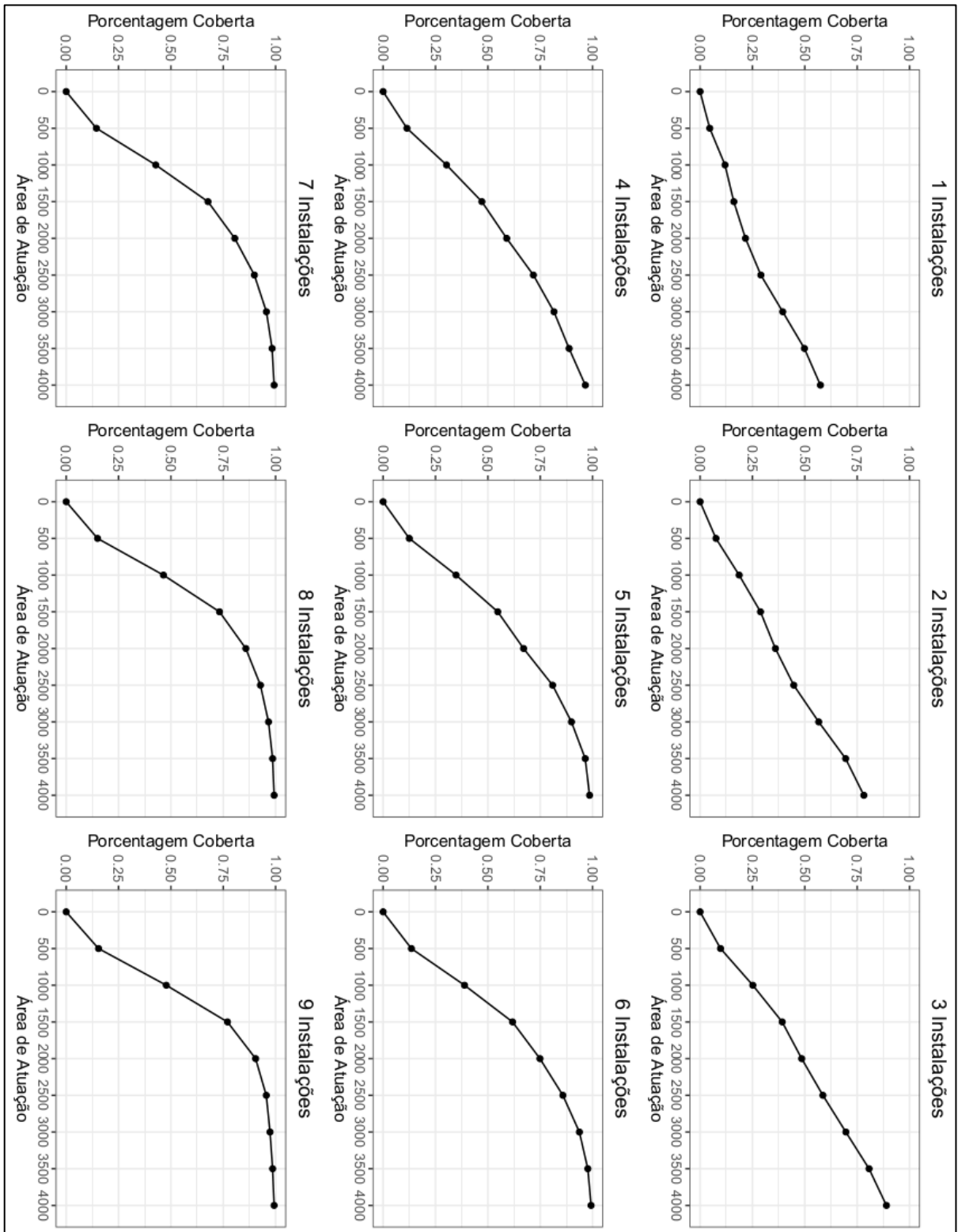
APÊNDICE B – ANÁLISE DE SENSIBILIDADE

Gráfico 17 – Análise com a quantidade de instalações fixas e raio variável



Fonte: O Autor (2019)

Gráfico 18 – Análise com o raio fixo e instalações variáveis



Fonte: O Autor (2019)