

Identificando os melhores locais para instalar câmeras de vigilância através da análise de dados abertos

Almir M. da Silva Neto¹, Iago M. da Conceição Silva¹,
João B. Rocha-Junior¹, Angelo A. Duarte¹

¹Universidade Estadual de Feira de Santana (UEFS)
Caixa Postal 292 – 294 – Feira de Santana – BA – Brasil

{almirneto338,iagomachadocs}@gmail.com, {joao,angeloduarte}@uefs.br

Abstract. *There is an increasing amount of open data available in the Internet. This data can be used to develop new useful tools for society. This paper presents a new tool that uses open data: news sites, OpenStreetMap and IBGE to indicate the best places to install new surveillance cameras in a city, taking into account the areas with higher crime rate, the favored population and the location of the current police stations.*

Resumo. *Existe uma quantidade cada vez maior de dados abertos disponíveis na Internet. Estes dados podem ser utilizados para desenvolver novas ferramentas úteis para a sociedade. Este artigo apresenta uma nova ferramenta que utiliza dados abertos: sites de notícias, OpenStreetMap e IBGE, para indicar os melhores locais para instalar câmeras de vigilância em uma cidade, levando em consideração as áreas com maior criminalidade, o tamanho da população atendida e a localização atual dos postos policiais.*

1. Introdução

Muitas cidades brasileiras possuem um alto índice de criminalidade. De acordo com dados do Atlas da Violência 2018 [Cerqueira et al. 2018], cerca de 553 mil pessoas morreram devido à violência intencional no Brasil nos últimos 10 anos. Independente dos motivos que levam uma cidade a ter essa característica, seus governantes, secretarias de segurança pública e a própria polícia procuram métodos para redução desses índices e aumento da segurança da sociedade.

Uma das soluções para reduzir a taxa de criminalidade de um bairro é a instalação de câmeras de vigilância. Como descrito por Neto (2010), esta se apresenta como uma medida efetiva para ampliar o alcance da vigilância realizada pela polícia, reprimindo a ação de indivíduos mal intencionados.

Este artigo apresenta a ferramenta EAGLE (*Enhanced Application for Geolocation Eyes*). Essa ferramenta utiliza dados sobre a criminalidade dos bairros de uma cidade, dados do IBGE e os estabelecimentos da cidade para computar o melhor local para a instalação de uma câmera de vigilância, baseando-se na taxa de criminalidade do bairro, na quantidade de pessoas e estabelecimentos.

A estrutura do artigo é composta por cinco seções. Na Seção 2 são descritas as fontes de dados abertos utilizadas e os dados que foram coletados de cada fonte. A Seção 3 retrata os métodos empregados no desenvolvimento da ferramenta, enquanto a Seção 4 explica o funcionamento da ferramenta desenvolvida.

2. Fontes de dados abertos

Neste artigo, optou-se por estimar o índice de criminalidade em uma dada região (bairro ou cidade) através da utilização de fontes de dados abertos IBGE, OpenStreetMap, sites de notícias e sites de estabelecimentos. Esta seção apresenta estas fontes de dados, indicando as informações que foram extraídas de cada uma delas.

2.1. IBGE

Para a realização da estimativa populacional são utilizados dados da estimativa do IBGE para a população residente no Brasil, Unidades da Federação e Municípios, com data de referência em 1º de julho de 2018. O IBGE disponibiliza dados populacionais para estados e municípios, mas não a nível de bairros.

Para extrair os dados foi desenvolvido um programa que lê o arquivo de entrada no formato .xls obtido no site do IBGE¹ e extrai as informações e as armazena em um arquivo no formato csv. Cada linha do arquivo gerado corresponde a uma cidade e contém as seguintes informações: estado, código, nome e população da cidade.

2.2. OpenStreetMap

O OpenStreetMap (OSM) é um mapa colaborativo que contém dados geográficos. Através do OSM é possível selecionar uma área no mapa e exportar um arquivo que contém elementos como *nodes*, que representam os pontos do mapa, *ways* que representam os percursos entre esses pontos e *relations*, que representam contornos de países, estados, cidades e bairros.

Além disso, o OSM disponibiliza informações a respeito dos elementos do mapa através de *tags*. Essas *tags* são conjuntos de chave e valor que descrevem características de algum elemento, como *node*, *way* ou *relation*. Através das *tags* é possível, por exemplo, reconhecer quando um percurso (*way*) é uma rodovia, verificando se este possui uma *tag* cuja chave seja *highway*, e saber quando uma rodovia é residencial, verificando se o valor desta *tag* é *residential* ou *living_street*.

Mais de 16GB de dados geográficos do Brasil estão disponíveis no OpenStreetMap. Contudo, esses dados estão condensados em um único arquivo, dificultando recuperar uma informação de uma região específica. Para simplificar esse processo, o arquivo disponibilizado pelo OSM é separado em arquivos menores, onde cada arquivo contém os dados de uma cidade do Brasil. Os polígonos das cidades são obtidos, utilizando o Nominatim². Cada polígono é um conjunto de pontos que formam o contorno da cidade.

Para tal separação, utiliza-se uma aplicação chamada Osmosis³, com ela é possível filtrar o arquivo do OSM, passando como parâmetro o arquivo obtido do OSM, o polígono da região de interesse e o diretório do arquivo de saída no formato OSM que receberá todas as informações referentes à região de interesse.

2.3. Sites de Notícias

A ferramenta utiliza dados sobre criminalidade acerca de um bairro com o objetivo de identificar o bairro com a maior taxa de criminalidade. Para isso, utiliza-se dados co-

¹ftp://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2018/estimativa_dou_2018_20181019.xls

²<https://nominatim.openstreetmap.org/>

³<https://wiki.openstreetmap.org/wiki/Osmosis>

letados do site de notícias local da cidade (ex: Acorda Cidade) que se deseja aplicar o algoritmo, preferencialmente da seção policial da mesma, verificando e contabilizando a quantidade de notícias em cada bairro da cidade.

Comumente, os sites de notícias possuem o mesmo padrão de formatação de suas páginas, fazendo com que seja possível coletar e armazenar os dados necessários a partir um robô. Caso o site possua uma seção policial recomenda-se que o dados sejam coletados a partir da mesma.

2.4. Estabelecimentos

Para obter os dados dos estabelecimentos foi utilizado um Robô [Fonseca 2018] que varre sites como GuiaMais, TripAdvisor e Yelp e extrai informações dos estabelecimentos como: endereço, telefone, categoria e nome do estabelecimento. Para facilitar a visualização destes estabelecimentos no mapa, fez-se necessário obter as coordenadas geográficas (latitude e longitude) dos mesmos. Estas coordenadas foram obtidas utilizando a API HERE Geocoder⁴. A base de dados de estabelecimentos contém mais de 900 mil estabelecimentos do Brasil.

3. Métodos

Alguns métodos são utilizados no tratamento dos dados coletados para o desenvolvimento e aplicação da ferramenta. Esses métodos são descritos nesta seção.

3.1. Obtendo automaticamente o contorno dos bairros

Para algumas cidades, o OSM ainda não tem dados sobre o contorno de bairros. Assim, para estas cidades, faz-se necessário criar um método que permita obter esses contornos de forma automática (regionalização). Desse modo, dois métodos são adotados para a regionalização do mapa na ferramenta desenvolvida. Um deles é utilizar o contorno dos bairros coletados do Nominatim para cada cidade.

O outro método criado para obter o contorno dos bairros se baseia no diagrama de Voronoi. O diagrama de Voronoi [Fortune 1995] de um conjunto de pontos (núcleo) é uma coleção de regiões que dividem um plano. Cada região corresponde a um núcleo, e todos os pontos em uma região são mais próximos deste núcleo que qualquer outro. Neste caso o conjunto de pontos utilizado é formado pelos pontos que correspondem a cada bairro da cidade, coletados do OSM.

3.2. Estimando a população de uma região

A estimativa populacional em cada região é feita com base no método utilizado por Bast et al.(2014), o qual assume que a população de determinada área está fortemente relacionada à soma do comprimento das ruas localizadas nessa área, uma vez que toda casa está próxima a alguma rua.

O cálculo do comprimento das ruas em cada região é feito utilizando um grafo, que guarda os pontos e percursos do mapa obtidos através do OpenStreetMap. Cada percurso das rodovias residenciais, identificadas como descrito na Seção 2.2, é armazenado no

⁴<https://developer.here.com/signup/geocoding>



Figura 1. Ruas desconectadas em uma célula

grafo como uma aresta entre os dois vértices nas suas extremidades. Apenas as arestas presentes dentro do contorno da cidade são inseridas no grafo.

Em cada aresta do grafo também são armazenados os comprimentos de cada percurso das rodovias, calculados através da fórmula de Haversine [Chopde and Nichat 2013], que utiliza a latitude e longitude dos vértices para calcular a distância entre eles.

Inicialmente pensou-se em calcular o somatório do comprimento das ruas em uma região partindo de um ponto da região e percorrendo todos os percursos da região com um algoritmo para percorrer o grafo em profundidade [Tarjan 1972]. A Figura 2 contém um trecho de um mapa de uma cidade com o contorno de de uma região do mapa. Os pontos marcados com um círculo preto indicam o início e o fim das ruas que já foram percorridas pelo algoritmo. Como não há uma conexão entre todas as ruas presentes na célula, o algoritmo não consegue visitar todas as ruas partindo de um ponto qualquer do grafo, pois as ruas de uma região formam um grafo desconexo.

Optou-se então por realizar o somatório do comprimento das ruas dentro de cada região através de uma iteração sobre todas as arestas do grafo, Dessa maneira garante-se que todas as arestas que pertencem a uma região serão consideradas. Caso ambos os vértices da aresta estejam dentro da região todo o comprimento da aresta é atribuído à região, caso apenas um dos vértices esteja dentro da região apenas metade do comprimento da aresta é atribuído.

A estimativa populacional realizada para uma região é proporcional à soma do comprimento das ruas residenciais presentes nessa região em relação ao comprimento total de ruas residenciais da cidade. Dessa forma, encontra-se a porcentagem do comprimento de ruas residenciais da cidade presentes na região e aplica-se essa porcentagem sobre a população total da cidade, chegando a uma estimativa populacional da região.

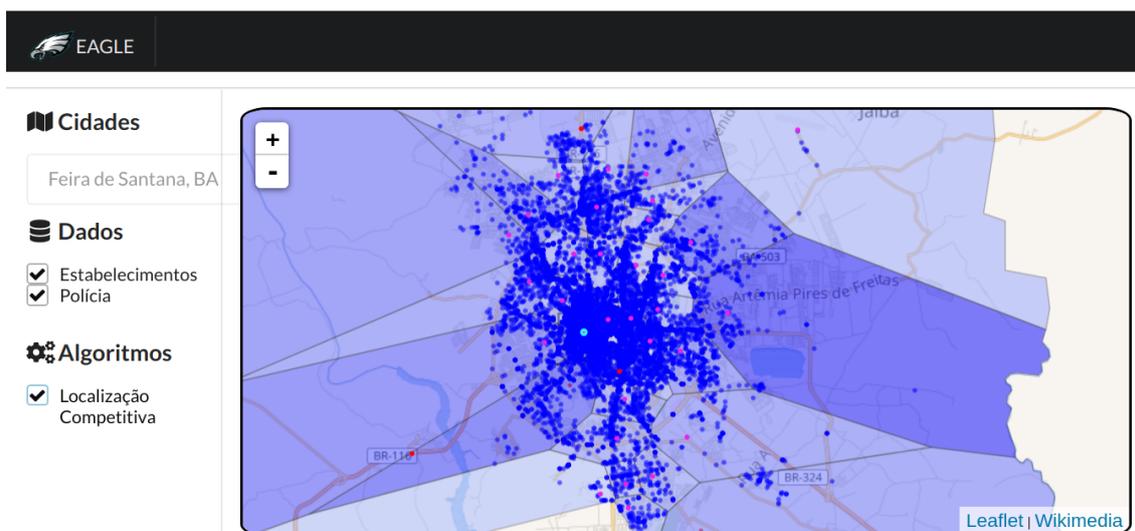


Figura 2. Tela Principal do Sistema

3.3. Identificando os melhores locais para a instalação de câmeras de vigilância

Para obter as melhores regiões para instalar as câmeras de vigilância, utilizou-se o Algoritmo de Localização Competitiva [Xiao et al. 2011]. Este algoritmo utiliza três conjuntos de pontos para encontrar o melhor local. As *facilities* são os bairros que já possuem postos policiais, os *clientes* são os estabelecimentos beneficiados pela instalação de uma câmera de vigilância e os locais candidatos são os possíveis pontos a receber uma câmera. A localização competitiva identifica qual o candidato que atrai uma maior quantidade de clientes em relação as *facilities* já existentes. Cada cliente atraído por uma *facility* é o cliente mais próximo dessa *facility* em comparação as outras. Ao fim do algoritmo teremos um conjunto de estabelecimentos atraídos por cada local candidato.

O melhor bairro para a instalação de uma câmera de vigilância é dado pela fórmula abaixo, onde e é a quantidade de estabelecimentos atraídos pelo bairro sobre a quantidade total de estabelecimentos, t é a taxa de criminalidade de um bairro dada pela proporção de notícias pela população estimada no bairro e α é um valor entre 0 e 1 inserido pelo usuário que define os pesos a serem levados em consideração para o cálculo do escore.

$$score = e * \alpha + (1 - \alpha) * t$$

O bairro que possuir o maior valor de **escore** é o melhor bairro para instalação da câmera de vigilância.

4. Aplicação

EAGLE é uma ferramenta web que permite aos usuários visualizar áreas de maior ocorrência de criminalidade de acordo com os dados coletados e realiza a escolha do melhor local para a instalação de câmeras de vigilância.

Ao abrir a ferramenta, o usuário pode escolher uma das cidades suportadas⁵ para a visualização das regiões. Ao clicar sobre uma região da cidade escolhida, são exibidas informações como a estimativa populacional calculada.

⁵Atualmente, Feira de Santana

O usuário também pode visualizar os estabelecimentos e os postos policiais. Cada estabelecimento e posto policial é representado por um ponto no mapa, como mostra a Figura 2, sendo os mesmos azuis e vermelhos, respectivamente. Ao clicar sobre um ponto que representa um estabelecimento, o usuário poderá ver as informações, como endereço, nome, telefone e categoria, do estabelecimento clicado.

Ao selecionar a *checkbox* Localização Competitiva na Seção Algoritmos, o mapa exibirá os pontos candidatos (rosas) e o melhor ponto (verde) para instalar uma câmera de segurança.

Os dados de cada região de cada cidade exibidos pela ferramenta são pré-processados e armazenados em arquivos, de modo que quando um dado é solicitado pelo usuário a ferramenta apenas realiza uma busca nos arquivos, evitando que os dados sejam gerados sempre que solicitados.

5. Conclusão

A ferramenta EAGLE mostra como dados coletados das fontes abertas, como o IBGE, OpenStreetMap e sites de notícias, podem auxiliar em questões de segurança pública como a identificação de áreas de maior ocorrência de criminalidade e escolha do melhor local para a instalação de câmeras de vigilância. Dados reais foram utilizados em seu desenvolvimento e implantação, permitindo que a ferramenta seja aplicada na prática contribuindo para a melhoria da segurança pública.

Referências

- Bast, H., Sternisko, J., and Storandt, S. (2014). Forestmaps: A computational model and visualization for forest utilization. In *International Symposium on Web and Wireless Geographical Information Systems*, pages 115–133. Springer.
- Cerqueira, D., Lima, R. S., Bueno, S., Neme, C., Ferreira, H., Coelho, D., Alves, P. P., Pinheiro, M., Astolfi, R., Marques, D., et al. (2018). Atlas da violência 2018. Instituto de Pesquisa Econômica Aplicada (IPEA).
- Chopde, N. R. and Nichat, M. (2013). Landmark based shortest path detection by using a* and haversine formula. *International Journal of Innovative Research in Computer and Communication Engineering*, 1(2):298–302.
- Fonseca, F. L. (2018). Recuperação e Extração Automática de Dados da Web. Trabalho de Conclusão do Curso de Engenharia de Computação, UEFS, Feira de Santana, Brasil.
- Fortune, S. (1995). Voronoi diagrams and delaunay triangulations. In *Computing in Euclidean geometry*, pages 225–265. World Scientific.
- Neto, A. D. C. (2010). O sistema de videomonitorização como ferramenta de policiamento preventivo. Centro de Aperfeiçoamento e Estudos Superiores, Polícia Militar do Estado de São Paulo.
- Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM J. Comput.*, 1(2):146–160.
- Xiao, X., Yao, B., and Li, F. (2011). Optimal location queries in road network databases. In *2011 IEEE 27th International Conference on Data Engineering*, pages 804–815. IEEE.