

# Simpósio de Integração Acadêmica

## “Ciências Básicas para o Desenvolvimento Sustentável”

SIA UFV 2023



## ELABORAÇÃO DE ROTAS SEGURAS PARA VEÍCULOS COLETIVOS URBANOS

Modalidade: Pesquisa | Área Temática: Ciência da Computação | Grande Área: Ciências Exatas e Tecnológicas

Miguel Antônio Ribeiro e Silva<sup>1</sup>, Gabriel Vitor da Fonseca Miranda<sup>1</sup>, Thais Regina de Moura Braga Silva<sup>1</sup>

<sup>1</sup>Laboratório de Inteligência em Sistemas Pervasivos e Distribuídos (NESPED-Lab)  
Instituto de Ciências Exatas e Tecnológicas (IEF).  
Universidade Federal de Viçosa (UFV) - Florestal – MG – Brasil  
{miguel.a.silva, gabriel.v.miranda, thais.braga}@ufv.br

### Introdução

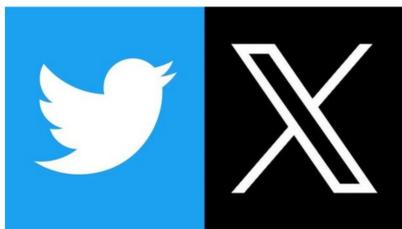
Grandes metrópoles enfrentam dilemas relacionados à criminalidade que impactam diretamente a vida de seus habitantes. É nesse contexto que exploramos, neste trabalho, as dinâmicas da segurança em ambientes urbanos, com um foco particular na cidade de São Paulo, buscando entender os desafios enfrentados em relação à criminalidade e à segurança nas rotas utilizadas por seus veículos coletivos.

### Objetivos

Buscamos identificar as áreas com maior incidência de crimes na cidade de São Paulo, com base na análise de postagens da rede social Twitter/X relacionados a eventos criminais.

Assim, os principais objetivos específicos foram:

1. Avaliar quais seriam as melhores palavras-chave para a filtragem das postagens relacionadas ao relato da ocorrência de crimes
2. Elaborar estratégias para a identificação da localização de um crime relatado por meio de uma postagem



### Material e Método

A metodologia utilizada na pesquisa envolveu o uso da linguagem Python, do ambiente Jupyter Notebook e da API v2 do Twitter.

1. Leitura de Artigos Relacionados: Foram realizados estudos de artigos relacionados a relatos de crimes em redes sociais com o objetivo de identificar o vocabulário utilizado pelos usuários ao relatar um crime. Isso ajudou a criar uma lista abrangente de termos e frases relacionados a crimes.
2. Testes de Busca na APIv2: Diversos testes de busca foram realizados na API do Twitter. Isso incluiu o uso das palavras-chave identificadas e suas variações verbais geradas pelo ChatGPT.
3. Estudo de Geolocalização: Isso incluiu a investigação de diferentes métodos de localização, como coordenadas geográficas, lugares (bairros, cidades, etc.), caixa delimitadora (bounding box), IDs associados a lugares e suas limitações.
4. Identificação de Locais no Texto dos *Tweets*: Também foram realizados estudos baseados em artigos para identificar locais mencionados no texto dos *tweets*. Isso envolveu o uso de técnicas de Processamento de Linguagem Natural (PLN) para extrair informações de localização de forma automatizada a partir do conteúdo dos *tweets*.

### Resultados e Discussão

1. Escolha de palavras-chave relevantes: No início da pesquisa, selecionamos palavras-chave como "furto", "assalto", "roubo" e "assédio" e suas variações verbais para encontrar *tweets* relacionados a crimes, descartando outras menos relevantes.
2. Mudança na obtenção de coordenadas geográficas: A capacidade de obter coordenadas geográficas diretamente dos *tweets* foi descontinuada em 2014, tornando-se disponível apenas para *tweets* que mencionam locais de outras redes sociais.
3. Desafios na associação com bairros em São Paulo: Após a descontinuação das coordenadas, a maioria dos *tweets* de São Paulo não estava associada a bairros específicos. Apenas uma pequena proporção dos 11.430 *tweets* analisados (1/145) continham informações de bairros, exigindo abordagens de Processamento de Linguagem Natural (PLN) para identificar locais nos textos.
4. Identificação de locais nos *tweets*: Para identificar locais mencionados nos *tweets*, revisamos 16 artigos acadêmicos que utilizaram técnicas como *Deep Learning* (CNN, BiLSTM, LSTM) e algoritmos como *Random Forest*, *Naive Bayes* e *Decision Tree*. Estamos avaliando a aplicabilidade dessas técnicas e considerando o uso de modelos pré-treinados, como BERT, para melhorar a precisão na identificação de locais.

### Conclusões

Este projeto busca estabelecer rotas seguras para veículos coletivos urbanos, usando dados de criminalidade municipal. Enfrentamos desafios na coleta e análise de *tweets* relacionados a crimes, principalmente na precisão da localização. Continuamos a pesquisa, explorando soluções baseadas em Processamento de Linguagem Natural, incluindo modelos como ChatGPT e BERT, para aprimorar a extração de informações de localização com maior precisão.

### Bibliografia

1. ALMEIDA, V. G. J. ; SILVA, THAIS R. M. BRAGA ; SILVA, FABRÍCIO A. . Se for, vá na paz: Construindo Rotas Seguras para Veículos Coletivos Urbanos. In: Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, 2022, Fortaleza. XL Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC), 2022.
2. MIRANDA, G. V. F. ; ALMEIDA, V. G. J. ; SILVA, THAIS R.M.B. ; SILVA, Fabrício Aguiar . Extração e Avaliação de uma Base de Dados sobre Criminalidade em Português a partir do Twitter. In: Simpósio Brasileiro de Computação Ubíqua e Pervasiva, 2023, João Pessoa. 15o. Simpósio Brasileiro de Computação Ubíqua e Pervasiva (SBCUP 2023), 2023.

### Agradecimentos

