EXPLORANDO O APRENDIZADO FEDERADO NA PREDIÇÃO DE VEÍCULOS ROUBADOS EM CIDADES INTELIGENTES

Guilherme S. Pires Veiga [1], Geraldo P. Rocha Filho [2]

RESUMO

Com a evolução das táticas de roubo de veículos, torna-se essencial explorar tecnologias de rastreamento mais robustas. Diante disso, este estudo explora o uso do aprendizado federado (FL) para estimar a localização de veículos roubados. O FL possibilita o treinamento de modelos diretamente em dispositivos locais, tais como sensores e câmeras, sem a necessidade de enviar dados a um servidor. Inserido no contexto de cidades inteligentes, este estudo concilia o desempenho computacional com a confidencialidade das informações, evitando a centralização e o risco de exposição de dados críticos. Os testes com um conjunto de dados reais avaliaram o modelo e comprovaram sua eficácia, demonstrando que as predições seguiram com uma alta precisão a localização do veículo ao longo do tempo.

PALAVRAS-CHAVE: aprendizado federado, privacidade de dados, rastreamento de veículos, cidades inteligentes.

EXPLORING FEDERATED LEARNING IN THE PREDICTION OF STOLEN VEHICLES IN SMART CITIES

ABSTRACT

As vehicle theft tactics evolve, it becomes essential to explore more robust tracking technologies. Therefore, this study explores the use of federated learning (FL) to estimate the location of stolen vehicles. FL allows models to be trained directly on local devices, such as sensors and cameras, without the need to send data to a server. Within the context of smart cities, this study balances computational performance with information confidentiality, avoiding centralization and the risk of exposing critical data. Tests with a real dataset evaluated the model and confirmed its effectiveness, demonstrating that predictions tracked the vehicle's location with high accuracy over time.

KEYWORDS: federated learning, data privacy, vehicle tracking, smart cities.

INTRODUÇÃO

O roubo de veículos é um desafio global de segurança, com milhões de ocorrências anuais, especialmente em áreas metropolitanas onde a criminalidade é mais acentuada

[1] Universidade Estadual do Sudoeste da Bahia - UESB, 202011757@uesb.edu.br
[2] Universidade Estadual do Sudoeste da Bahia - UESB, geraldo.rocha@uesb.edu.br

(Kairouz et al., 2021; NICB, 2023). A dificuldade na recuperação é agravada pela rápida realocação ou desmanche dos automóveis para o mercado de peças, o que exige o desenvolvimento de soluções tecnológicas mais eficientes (Interpol, 2022; CARFAX, 2023).

A evolução de tecnologias como a Internet das Coisas (IoT) e a Inteligência Artificial (IA) tem permitido novas estratégias de combate, utilizando sensores e análise de dados para identificar atividades suspeitas e acelerar a recuperação de ativos (Avnet, 2023). Dentre elas, o aprendizado federado (FL) se sobressai por sua abordagem distribuída, que preserva a privacidade ao permitir que dispositivos, como câmeras e sensores veiculares, treinem modelos de forma colaborativa sem centralizar ou compartilhar dados sensíveis.

O FL pode ser aplicado para antecipar as rotas de veículos roubados, desenvolvendo um modelo preditivo robusto que não apenas auxilia na recuperação, mas também garante a confidencialidade dos dados. Em cidades inteligentes, essa técnica permite que sensores e câmeras colaborem no treinamento de modelos de machine learning sem troca direta de informações (Kairouz et al., 2021). Ao descentralizar o processamento, o FL melhora o tempo de resposta em situações críticas, como roubos, e sua utilidade se expande para otimização de tráfego e rotas, firmando-se como uma solução chave para a gestão e segurança do transporte (Li et al., 2022; McMahan et al., 2020).

Nesse contexto, este trabalho investiga o uso do FL em sistemas de transporte inteligente para prever a localização de veículos com alerta de roubo. A abordagem utiliza o algoritmo Random Forest (RF) para refinar a predição de rotas. O objetivo é aprimorar a segurança por meio de uma previsão mais precisa dos trajetos de veículos roubados, apoiando a formulação de políticas urbanas baseadas em dados e fomentando um ambiente mais seguro nas grandes cidades.

MATERIAIS E MÉTODOS

Neste trabalho, foi empregada uma abordagem de FL para o rastreamento descentralizado da localização de veículos com alerta de roubo. A tarefa de predição foi conduzida pelo algoritmo RF, selecionado por sua alta capacidade de processar grandes conjuntos de dados e múltiplas variáveis, o que assegura predições robustas e

precisas. A estrutura federada foi fundamental para permitir que os modelos fossem treinados em dispositivos distribuídos sem a partilha direta de dados brutos.

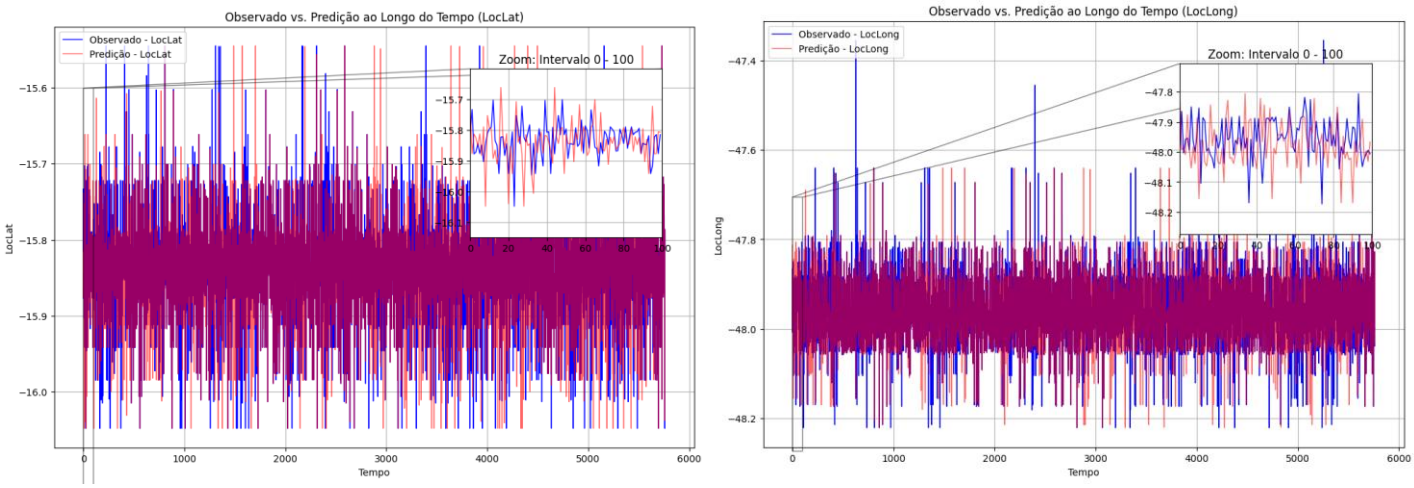
O conjunto de dados utilizado consistiu em registros anônimos de veículos roubados, com informações detalhadas sobre localização e os alertas emitidos. Todo o volume de dados foi concentrado no Distrito Federal, que serviu como cenário principal para o experimento. Para simular o ambiente federado, esses dados foram particionados e distribuídos entre os clientes (dispositivos), que treinaram os modelos usando apenas suas porções de dados locais, sem qualquer intercâmbio de informações entre si.

A precisão preditiva do modelo foi o foco da avaliação. Para isso, a performance foi mensurada tanto pela análise visual dos gráficos de predição (Observado vs. Predito) quanto pelo cálculo de métricas estatísticas de erro, como o Erro Médio Absoluto (MAE), Erro Quadrático Médio (MSE) e a Raiz do Erro Quadrático Médio (RMSE). Este modelo foi obtido ao final de 10 rodadas de um processo de treinamento federado, onde o servidor central agregava as atualizações dos modelos treinados localmente nos dispositivos clientes para gerar um modelo global a cada rodada.

A implementação do projeto foi realizada na linguagem Python. A estrutura de aprendizado federado foi construída com a biblioteca Flower (versão 2.0.0), enquanto a modelagem foi apoiada pelo PyTorch (versão 2.4.1). Para a manipulação e o processamento de dados, foram utilizadas as bibliotecas NumPy (versão 2.1.1) e Pandas (versão 2.2.3), e os gráficos de resultados foram gerados com o Matplotlib (versão 3.9.2).

RESULTADOS E DISCUSSÃO

A avaliação do desempenho do modelo foi focada em sua capacidade preditiva. A Figura 1 ilustra a comparação entre os dados de localização reais do veículo (linha azul) e os valores gerados pelo modelo (linha vermelha), permitindo uma análise qualitativa da sua precisão.



(a) Latitude (b) Longitude

Figura 1 - Comparativo entre valores reais (observado) e preditos pelo modelo para as coordenadas de (a) Latitude e (b) Longitude.

Na Figura 1a, que apresenta o desempenho para a latitude, observa-se uma notável aderência entre as duas linhas. Isso indica que o modelo foi capaz de acompanhar as variações e a tendência dos dados reais com alta precisão, o que valida a eficácia do algoritmo em aprender os padrões de deslocamento. De forma análoga, a Figura 1b detalha os resultados para a longitude, onde o comportamento consistente se repete e a linha de predição segue a trajetória observada com grande fidelidade. Essa consistência em ambas as coordenadas é crucial, pois demonstra a capacidade do modelo em reconstruir a localização bidimensional do veículo de maneira coesa.

Essa acurácia visual é confirmada e quantificada pelas métricas de erro da Tabela 1. O MAE demonstra que a média de erro das previsões é muito baixa, sendo de 0.056 graus para latitude e 0.061 para longitude. O MSE, que penaliza erros maiores, e sua raiz, o RMSE, corroboram essa análise. Os valores do RMSE (0.080 e 0.091) se mantêm próximos ao MAE, indicando que o modelo não apenas acerta na média, mas também não comete grandes erros esporádicos. Os baixos valores consistentes nessas três métricas atestam a alta precisão e a confiabilidade do modelo.

Métricas	Latitude (LocLat)	Longitude (LocLong)
MAE	0.056	0.061
MSE	0.006	0.008
RMSE	0.080	0.091

Em conjunto, a análise visual e quantitativa reforça a robustez da abordagem, validando a solução como uma ferramenta de rastreamento eficiente que respeita a privacidade dos dados.

CONCLUSÕES

Este trabalho validou a eficácia do FL na predição de rotas de veículos com alerta de roubo, utilizando uma abordagem que garante a privacidade das informações. A arquitetura descentralizada do FL provou ser eficiente para distribuir o processamento computacional entre os dispositivos clientes, confirmando a viabilidade da solução para aplicações de segurança pública.

A análise dos resultados confirmou a alta performance do modelo de predição. A avaliação qualitativa, baseada nos gráficos comparativos, demonstrou uma notável aderência entre as localizações previstas e os dados reais ao longo do tempo. Essa precisão foi corroborada pela análise quantitativa, cujas métricas de erro (MAE, MSE e RMSE) apresentaram valores baixos, validando o sistema como uma ferramenta precisa e confiável para o rastreamento de veículos.

Como trabalhos futuros, sugere-se investigar a robustez do sistema contra ataques adversariais, uma preocupação de segurança em ambientes federados.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., et al. (2021). Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*.
2. NICB (National Insurance Crime Bureau). *Vehicle Theft Statistics 2023*.
3. Interpol (2022). *International operation targets trafficking of stolen motor vehicles*.
4. CARFAX for Police (2023). *Vehicle Theft Exporting*.
5. Avnet Silica. *IoT in Stolen Vehicle Recovery: It's a Steal!*.
6. Li, T., Sahu, A. K., Talwalkar, A., Smith, V. (2022). Federated Learning: Challenges, Methods, and Future Directions. *IEEE Signal Processing Magazine*.

7. McMahan, B., Ramage, D., Talwar, K., y Arcas, B. A. (2020). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics*.