**UNIVERSIDADE DE PERNAMBUCO**

 

**Escola Politécnica de Pernambuco**

**Coordenação Setorial de Pós-graduação e Pesquisa**

**MARCOS AURÉLIO SILVA DE SOUZA**

**Dependability, Data Diversity and Deep Learning para detecção de FDIA**

Projeto de Pesquisa apresentado ao programa de Pós-graduação *Lato Sensu* em Inteligência Artificial, da Escola Politécnica de Pernambuco da Universidade de Pernambuco, como parte dos requisitos para a obtenção do título de Especialista em Inteligência Artificial.

Recife-PE, junho de 2021

**SUMÁRIO**

[**1. INTRODUÇÃO** 3](#_Toc76131643)

[**1.1.** **Motivação** 3](#_Toc76131644)

[**1.2.** **Apresentação do Problema** 4](#_Toc76131645)

[**1.3.** **Técnicas e ferramentas** 5](#_Toc76131646)

[**2. OBJETIVOS** 6](#_Toc76131647)

[**2.1** **Objetivos Gerais** 6](#_Toc76131648)

[**2.2 Objetivos Específicos** 6](#_Toc76131649)

[**3. METODOLOGIA** 7](#_Toc76131650)

[**4. CRONOGRAMA** 9](#_Toc76131651)

[**5. RESULTADOS ESPERADOS** 10](#_Toc76131652)

[**6. REFERÊNCIAS BIBLIOGRÁFICAS** 11](#_Toc76131653)

# **1. INTRODUÇÃO**

No mundo cibernético devido ao rápido crescimento da Internet e dos sistemas adaptativos complexos, como nos sistemas que compõe os veículos autônomos, como um sistema de controle adaptativo de navegação ou direção, o ataque de injeção falsa de dados (FDIA) [10] compromete a segurança desses sistemas. Essas invasões estão também interessado em ataques semelhantes em outros domínios de aplicações como saúde [12], finanças, defesa e energia. Isso torna o FDIA uma das questões prioritárias.

Como evitar falhas de segurança de funcionamento ou possibilidade de invasão de sistema de navegação ou direção de carros autônomos?

Esse trabalho busca fazer um sumário da intersecção dos seguintes assuntos: Dependability [22], Data Diversity [23] e Deep Learning [21], se propondo a sugerir caminhos para a junção desses assuntos, no estudo de caso para sistemas adaptativos de navegação de veículos autônomos.

* 1. **Motivação**

A aplicação da técnica de Data Diversity usando mecanismos de tomada de decisão onde o conjunto de dados de entrada expressem a possibilidade de detecção de dados inválidos e a busca em levar o sistema para um estado seguro, impossibilitando ou dificultando ataques de FDIA. Em um cenário de um desses ataques cibernéticos em um sistema de veículo autônomo, em que os dados dos sensores sendo manipulados para tomar uma decisão de uma frenagem indevida, o que ocorreria? Poderia acarretar em um acidente, certamente, causando dano a vidas humanas.

O conceito de ataque de injeção falsa de dados (FDIA) foi introduzido originalmente no domínio da rede inteligente[10]. Embora o termo pareça comum, ele significa especificamente o caso em que um invasor compromete as leituras do sensor de maneira tão complicada que erros não detectados são introduzidos nos cálculos de variáveis ​​e valores de estado. Devido ao rápido crescimento da Internet e dos sistemas adaptativos complexos associados, os invasores cibernéticos estão interessados ​​em explorar ataques semelhantes em outros domínios de aplicação, como saúde, finanças, defesa, governança, etc.

No mundo cibernético cada vez mais perigoso de sistemas adaptativos complexos de hoje, FDIA tornou-se uma das questões prioritárias a serem tratadas. É uma necessidade hoje para uma maior consciência e melhor mecanismo para conter tal ataque no ciberespaço.

Propor uma técnica que seja uma junção de Dependability com uso de Inteligência Artificial (IA) [19] para detectar uma invasão ou falha dos dados oriundos dos sensores para que o sistema tome precauções ou decisões de como agir com a tentativa de invasão ou falha, é um fator motivador para esse trabalho.

* 1. **Apresentação do Problema**

Como evitar falhas de segurança de funcionamento ou prevenção de ataques de injeção falsa de dados ao sistema de navegação ou direção de veículos autônomos [3]?

Na segunda metade da última década se viu um surgimento significativo de veículos comercialmente disponíveis com capacidade de direção autônoma. Esses sistemas de direção autônoma dependem fortemente da qualidade dos dados dos sensores para tarefas como localização e navegação [2]. Nos casos de dados oriundos de sensores nos veículos autônomos que são usadas unidades de controle eletrônico (ECUs), onde os ECUs se comunicam entre si seguindo um protocolo de comunicação padrão. Se ocorre uma invasão na rede do veículo para comprometer os ECUS, existem modelos de Deep Learning [20] para proteger, mas estudos recentes mostram que os invasores podem facilmente enganar os modelos [1]. No caso dessas invasões corrompe dados de oriundos de sensores que são usados no ECUS podem comprometer os sistemas avançados de assistência ao motorista (ADAS) ou o controle de cruzeiro adaptativo (ACC).

Carros autônomos usam atualmente os sensores Lidar e Radar, além de câmeras. Focarmos especificamente nesse trabalho, nos sensores do tipo Lidar [15] [18].

Dois atributos não-funcionais importantes não são atendidos nos ataques FDIA nesses casos, a confiabilidade e a segurança [5] . Como se trata de sistemas críticos em tempo real, a necessidade de técnicas de Dependability [6], se faz mais do necessária para lidar com esse problema.

* 1. **Técnicas e ferramentas**

Na apresentação do problema citamos o uso de Dependability, e uma delas, é a diversidade de dados (Data Diversity) [23] é descrita e os resultados de um estudo piloto são apresentados. As regiões do espaço de entrada que causam falha para certos programas experimentais são discutidas e a reexpressão de dados, a maneira pela qual conjuntos de dados de entrada alternativos podem ser obtidos, é examinada. É fornecida uma descrição do bloco de nova tentativa, que é o equivalente com dados diversificados do bloco de recuperação, e um modelo do bloco de nova tentativa, juntamente com alguns resultados empíricos, é apresentado. A programação de N-copy [23], que é o equivalente de dados diversificados da programação de N-version, é considerada, e um modelo simples e alguns resultados empíricos também são fornecidos.

Na votação de aprovação, cada entrada para o processo de votação consiste em um conjunto finito ou infinito de valores que foram "aprovados" pelo canal de computação correspondente e o valor, ou conjunto de valores, com a votação de aprovação mais alta deve emergir como saída.

Para uso de conjuntos de dados válidos e detecção de dados inválidos podemos treinar modelos que expressaram esses conjuntos que serão usados na parte do algoritmo de votação da Data Diversity. Das técnicas de Deep Learning que propromos usar estão as de Redes Profundas Convolucionais (CNNs), especificamente a FT-CNN [7] e [Redes Adversariais Generativas](https://arxiv.org/abs/1406.2661) (GNAs) [17].

Portanto, este trabalho apresenta uma técnica para prevenção de FDIA e propor contra-medidas para possíveis ataques a sistemas de veículos autônomos [9] [14].

**2. OBJETIVOS**

* 1. **Objetivos Gerais**

O projeto visa identificar padrões de dados válidos e inválidos, prever falhas oriundas de dados usados nos sistemas de veículos autônomos, mas especificamente o de controle de direção ou navegabilidade, vindo de sensores, usando IA com técnicas de Dependability.

**2.2 Objetivos Específicos**

* Realizar um breve estudo sobre a Dependabiliy, Data Diversity e Deep Learning;
* Descobrir padrões de dados válidos;
* Descobrir padrões de dados inválidos;
* Detectar um ataque de injeção falsa de dados [13];
* Data Diversity [14] como uma premissa para tomar contramedidas oportunas para evitar a propagação da falhas ocasionadas por ataques *FDIA*;
* Técnicas de IA para reexpressar novamente os dados de entrada, para uso em técnicas de Diversidade de Dados;
* O estudo comparativo do de Redes Profundas Convolucionais (*CNNs*) e [Redes Adversariais Generativas](https://arxiv.org/abs/1406.2661) (*GNAs*) para reexpressão de conjuntos de dados de entrada.

**3. METODOLOGIA**

Nesse projeto, será utilizada uma base de dados livre para uso não-comercial,

“Mini”, disponível na Nuscenes [4], contendo metadata e dados dos sensores (lidar, radar and camera) e um tamanho total de 3.88 GB. Usaremos essa base para treinamento. Para teste, usaremos o dataset Test, com tamanho de 22.46 GB para sensores do tipo Lidar.

Outra base que podemos experimentar é uma expansão, a “CAN bus expansion”, disponível na Nuscenes [4]. A base em questão contém dados de baixo nível do veículo sobre a rota do veículo, IMU, pose, feedback do ângulo de direção, bateria, freios, posição da marcha, sinais, velocidades das rodas, acelerador, torque, sensores solares, odometria e muito mais.e um tamanho total de 0.73 GB.

O conjunto de dados do sensor Lidar consiste em:

**scene**: trecho de 25-45 segundos da jornada de um carro.

**sample**- Um instantâneo anotado de uma cena em um determinado carimbo de data / hora.

**sample**\_data - dados coletados de um sensor específico.

**sample\_annotation** - Uma instância anotada de um objeto de nosso interesse.

**instance** - Enumeração de todas as instâncias do objeto que observamos.

**category** - Taxonomia de categorias de objeto (por exemplo, veículo, humano). Normalmente denotado por um número

**atribute** - Propriedade de uma instância que pode mudar enquanto a categoria permanece a mesma.

**visibility** - Será usado na próxima versão

**sensor** - Um tipo de sensor específico.

**calibrated\_sensor** - Definição de um sensor específico conforme calibrado em um veículo específico.

**ego\_pose** - O veículo do ego posa em um determinado carimbo de data / hora.

**log**- Informações de log das quais os dados foram extraídos.

**map**- Mapear dados que são armazenados como máscaras semânticas binárias de uma visão de cima para baixo.

Já com a base:

* Devermos ter diversidade de dados descrita e os resultados de um estudo piloto apresentados.
* O uso de GAN (Generative adversarial networks) para expressar novamente os dados de entrada em técnicas de Diversidade de Dados [16][17][23].
* O uso de FT-CNN**:** (Fault Tolerance for Convolutional Neural Networks) para expressar novamente os dados de entrada em técnicas de Diversidade de Dados [7].
* Uso do GAN para gerar conjuntos de entrada de dados inválidos para treinamento do modelo do Votador [16][17] [21][23].
* Modelo de Votador do N-copy programing detecta possível FDIA [23].
* Provar que a randomização dos dados de entrada pode ajudar a impedir ataques cibernéticos, negando informações críticas sobre um sistema alvo conhecido anteriormente por um invasor. A randomização das características do software pode ajudar a impedir ataques cibernéticos, negando informações críticas sobre um sistema alvo conhecido anteriormente por um invasor. [5]

# **4. CRONOGRAMA**

01 – Pesquisa Bibliográfica;

02 – Implementação do algoritmo;

03 – Validação do modelo e técnica proposta nesse trabalho;

04 – Avaliação das métricas obtidas comparativamente em cada técnica de IA para reexpressão de dados;

05 – Levantamento de perspectivas futuras;

06 – Escrita do artigo.

As atividades serão executadas conforme o cronograma a seguir:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Cronograma | | | | | | |
| Atividade | 2021 | | | | | |
|  | Junho | Julho | Agosto | Setembro | Outubro | Novembro |
| 01 |  |  |  |  |  |  |
| 02 |  |  |  |  |  |  |
| 03 |  |  |  |  |  |  |
| 04 |  |  |  |  |  |  |
| 05 |  |  |  |  |  |  |
| 06 |  |  |  |  |  |  |

# **5. RESULTADOS ESPERADOS**

* Provar a confiabilidade alta do uso de Data Diversity com Deep Learning como técnica de detecção de FDIA.
* Sugerir contramedidas oportunas para evitar a propagação da falha em ataques FDIA [8].
* Trabalhos futuros em prevenção de ataques FDIA em dispositivos de saúde, energia e IOT [11][12]

# **6. REFERÊNCIAS BIBLIOGRÁFICAS**

[1] Li, Yi; Li, Jing e Xiong, Kaiqi. An Adversarial Attack Defending System for Securing In-Vehicle Networks. [2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)](https://ieeexplore.ieee.org/xpl/conhome/9369428/proceeding). 9-12 Jan. 2021

[2] Balakrishnan, Arjun; Rodriguez Flore, Sergio; e Reynaud ,Roger. Integrity Monitoring of Multimodal Perception System for Vehicle Localization. CNRS, ENS Paris-Saclay, Université Paris-Saclay, 91190 Gif-sur-Yvette, France.

[3] [Zinoune](https://www.researchgate.net/profile/Clement-Zinoune), [Clément](https://www.researchgate.net/profile/Clement-Zinoune); [Bonnifait](https://www.researchgate.net/profile/Philippe-Bonnifait), [Philippe](https://www.researchgate.net/profile/Philippe-Bonnifait); Ibanez-Guzman,[Javier](https://www.researchgate.net/profile/Javier-Ibanez-Guzman). Sequential FDIA for Autonomous Integrity Monitoring of Navigation Maps on Board Vehicles. IEEE Transactions on Intelligent Transportation Systems (IEEE T INTELL TRANSP)

[4] nuScenes: A multimodal dataset for autonomous driving <https://www.nuscenes.org/>

[5] Knight ,John C;. Davidson, Jack.;Tuong, Ann N.; Hiser, Jason; Co, Michele. Diversity in Cybersecurity. Publicado em Computer IEEE, ( Volume: 49, Issue: 4, Apr. 2016).

[6] Amman, Paul. Data Redundancy for the Detection and Tolerance of Software Faults. Publicado em Computing Science and Statistics, pp 43-52.

[7] Zhao, Kai; Di, Sheng; Li, Sihuan; Liang, Xin; Zhai, Yujiia; Chen, Jieyang; Ouyang, Kaiming; Cappello, Franck; Chen, Zizhong.FT-CNN: Algorithm-Based Fault Tolerance for Convolutional Neural Networks. Publicado em 31 de dezembro de 2020, IEEE Transactions on Parallel and Distributed Systems ( Volume: 32, Issue: 7, July 1 2021), Page(s):1677 – 1689.

[8] Monperrus, Martin. Automatic Software Repair: A Bibliography. Publicado em ACM Computing Surveys, Association for Computing Machinery, 2017, 51, pp.1-24

[9] Houben, C.; Houben, S. Endowing advanced driver assistance systems with fault tolerance. Publicado em Annual Reviews in Control, Volume 39, 2015, Pages 58-67.

[10] Ahmed, Mohiuddin; Pathan, Al-Sakib Khan. False data injection attack (FDIA): an overview and new metrics for fair evaluation of its countermeasure. Publicado em Complex Adapt Syst Model 8, 4 (2020).

[11] Ahmed, Mohiuddin; Bostami, Biozid; Choudhury, Salimur. False Data Injection Attacks in Internet of Things. Publicado no livo Performability in Internet of Things (pp.47-58), janeiro de 2019.

[12] Ahmed, Mohiuddin; Ullah, Abu S. S. M. Barkat.False Data Injection Attacks in Healthcare. Publicado em 15th Australasian Conference, AusDM 2017, Melbourne, VIC, Australia, August 19-20, 2017.

[13] Liu, Lanchao; Esmalifalak, Mohmmad; Ding, Qifeng; Emesih, Valentin A.; Han, Zhu. Detecting False Data Injection Attacks on Power Grid by Sparse Optimization. Publicado em IEEE TRANSACTIONS ON SMART GRID, VOL. 5, NO. 2, MARCH 2014.

[14] Nguyen-Tuong, Anh; Evans, David; Knight ,John C.; Cox, Benjamin; Davidson, Jack W. .Security through redundant data diversity. Publicado em International Conference on Dependable Systems & Networks: Anchorage, Alaska, June 24-27 2008.

[15] Zhang, Yihuan ; Wang, Liang; Jiang, Xuhui; Zeng ,Yong; e Dai, Yifan. Sensor Lidar An efficient LiDAR-based localization method for self-driving cars in dynamic environments. Publicado em Cambridge University Press, 20 April 2021.

[16] Malof, ordan M.; Reichman, Daniël; Collins, Leslie M. . Data Augmentation in feature-space with Generative Adversarial Networks, applied to GPR-based Buried Threat Detection. Publicado em 10th International Workshop on Advanced Ground Penetrating Radar, Sep 2019, Volume 2019, p.1 - 7

[17] Li, Dan; Chen, Dacheng; Jin, Baihong; Shi, Lei. MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks. Publicado em Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series (pp.703-716)

[18] Badue, Claudine; Guidolini, Rânik; C. V. Raphel; Azevedo, Pedro; Forechi, Avelino. Self-Driving Cars: A Survey. Publicado em Expert Systems with Applications Volume 165, 1 March 2021, 113816.

[19]Souza, Marcos. Linguagem ADA: estado da arte, tecnologias associadas e mecanismos de Dependabilidade. Projeto de conclusão de Curso de Bacharelado em Engenharia da Computação, 2011, Universidade de Pernambuco.

[20]Mode, Gautam Raj; Calyam, Prasad; Hoque, Khaza Anuarul. False Data Injection Attacks in Internet of Things and Deep Learning enabled Predictive Analytics. Publicado em IEEE NOMS 2020 conference.

[21] Su, Feng; Yuan, Peijiang; Wang, Yangzhen; Zhang, Chen. The superior fault tolerance of artificial neural network training with a fault/noise injection-based genetic algorithm. Publicado em Protein Cell 7, 735–748 (2016).

[22] Laprie, J. C. Dependable Computing: Concepts, Limits, Challenges. 25th IEEE Internacional Symposium on Fault-Tolerant Computing, Pasadena, California, USA, June 27-30, 1995, Special Issue, p. 42-54.

[23] Amman, Paul; Knight ,John C. Data Diversity: An Approach to Software Fault Tolerance. Publicado em IEEE TRANSACTIONS ON COMPUTERS, VOL. 31. NO. 4. APRIL 1988.