Planificación para el Desarrollo del Proyecto Integrador

Alex Pérez

11 de septiembre de 2024

Contenido

Título del Proyecto

Machine Learning Applied for Cybersecurity of Energy Management Systems

(Aprendizaje Automático Aplicado a la Ciberseguridad del Manejo de Sistemas Energéticos)

Tutor: Felipe Grijalva

Autor: Alex Pérez

Relevancia y Justificación

Este proyecto se enfoca en mejorar la seguridad y confiabilidad de los sistemas energéticos basados en paneles solares, desarrollando un modelo de redes neuronales recurrentes (RNNs) para detectar alteraciones en los datos de generación y consumo de energía. Utilizando técnicas de aprendizaje automático, el modelo identificará patrones anómalos en los datos reportados por transformadores, permitiendo detectar manipulaciones o fraudes. Esto aborda la creciente vulnerabilidad de los sistemas digitales de monitoreo energético a ciberataques, garantizando la integridad y eficiencia del suministro eléctrico.

Objetivos

Objetivo General:
Mejorar la precisión en la predicción en los datos de electricidad generada por paneles solares y detección de alteraciones en los mismos.
Objetivos Específicos:

- Desarrollar modelos avanzados para el análisis de mediciones no alteradas.
- Utilizar redes neuronales recurrentes para predecir la energía generada y demandada.
- Implementar técnicas de aumento de datos para simular variaciones en la generación de energía.

Metodología de Trabajo

Fase Experimental:

- **Herramientas:** Google Colab, Termius y una VPN para acceder a un servidor con mayor capacidad de cómputo.
- **Flujo de trabajo:** Conexión remota desde Google Colab a un servidor universitario vía port forwarding y Docker.

Modelos y Estructuras:

- **SISO (Single Input, Single Output):** Predicción independiente para cada transformador utilizando arquitecturas variadas.
- **MIMO (Multiple Input, Multiple Output):** Predicción conjunta para todos los transformadores.

Detección de Ataques FDIA y Análisis de Resultados

Modelos Entrenados:

- LSTM, Transformers, TCN, y un modelo híbrido LSTM+TCN para predecir demanda y generación.
- Datos: 25,000 puntos por transformador (17 transformadores). Los conjuntos de entrenamiento, validación, prueba y testeo serán divididos de manera acorde a los datos disponibles.

Detección de FDIA:

- **Metodología:** Introducción de perturbaciones simulando ataques FDIA y comparación de datos manipulados con predicciones.
- **Técnicas de Detección:** Distancia media y prueba de Chi-cuadrado, usando un vector de 34 características.

Análisis de Resultados:

- Evaluación del tiempo de entrenamiento y precisión de los modelos usando Wandb para análisis detallado.
- Comparación de modelos en términos de precisión y detección de ataques.

Resultados y entregables

- **Repositorio en GitHub**: Se proporcionará un repositorio con el código fuente de los modelos desarrollados, la configuración de los experimentos y los scripts de detección de FDIA. Este repositorio incluirá documentación detallada para replicar los experimentos, enlaces a los notebooks de Google Colab utilizados para el entrenamiento, así como un informe que describe los resultados, análisis realizados, conclusiones y recomendaciones para trabajos futuros.

Referencias

- Zhang, Y., Wang, X., & Liu, J. (2020). Detecting False Data Injection Attacks in Smart Grids Using CNNs and LSTMs. *IEEE Transactions on Smart Grid*, 11(4), 3043-3051. https://arxiv.org/pdf/2006.11477.
- Stanford University. (2024). Recurrent Neural Networks cheatsheet. Recuperado de `https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks`.
- Cayci, S., & Eryilmaz, A. (2024). Convergence of Gradient Descent for Recurrent Neural Networks: A Nonasymptotic Analysis. *arXiv preprint arXiv:2402.12241*. Recuperado de `https://arxiv.org/abs/2402.12241`.
- Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv preprint arXiv:1909.09586*. Recuperado de `https://arxiv.org/abs/1909.09586`.