

Síntesis informativa sobre actividades de investigación en el DIIT Resúmenes Didácticos de los trabajos originales

Utilidad: difundir ideas, metodologías y procedimientos que pueden ser de utilidad para obtener mayor aprovechamiento en las actividades y mejoras.

Finalidad: transmitir información sintética sobre actividades que se efectúan en cada proyecto

Aprendizaje automático para el control del estado de salud en sistemas Aeroespaciales

Código:C230

Director: Jorge Ierache

e-mail: jierache@unlam.edu.ar

Co-Director: Pablo Soligo

e-mail: psoligo@unlam.edu.ar

Integrantes:

Investigadores:

Alumnos Ad-Honorem:

Becarios: Germán Merkel

Síntesis del contenido

El proyecto C230 pretende desarrollar el potencial de las técnicas de minería de datos y aprendizaje automático aplicado a los sistemas del área aeroespacial. Tiene como objetivo encontrar patrones de comportamiento que puedan anticipar averías o descubrir automáticamente potenciales fallos (Soligo & Ierache, 2019). Actualmente el grupo de investigación GIDSA (Grupo de Investigación y Desarrollo Aeroespacial de la UNLaM) cuenta con prototipo funcional de segmento terreno (UGS) (www.ugs.unlam.edu.ar) y telemetría satelital donde implementar las soluciones exploradas.

Temario de presentación

El grupo de investigación GIDSA propone, mediante prototipos, aplicar a la industria aeroespacial soluciones de software de uso extendido y alta penetración en la industria informática de propósito general. La estrategia mayormente utilizada para el control del estado de salud de los satélites se basa en el control de límites (valores máximos y mínimos) sobre la telemetría (Yairi, y otros, 2004). Esta técnica, de sencilla implementación, no asigna valor a la historia ni al contexto. En el grupo de investigación nos proponemos explorar la factibilidad de implementar las hoy difundidas técnicas de minería de datos y aprendizaje automático como herramienta para detectar patrones anormales sobre la telemetría satelital.

Metodología del trabajo desarrollado

El grupo de investigación cuenta con telemetría satelital de tiempo real proveniente principalmente de la red satnogs (<https://satnogs.org/>) además de telemetría histórica de otras fuentes alternativas. Cuenta con un prototipo de segmento, UGS, (www.ugs.unlam.edu.ar) terreno capaz de decodificar, interpretar y almacenar de forma normalizada telemetría de varios satélites. Sobre los trabajos de (Yairi, Kawahara, Fujimaki, Sato, & Machida, 2006), (Yairi, Kawahara, Fujimaki, Sato, & Machida, 2006) y (Polson, 2019), aplicando parcialmente metodologías CRISP-DM (Proceso estándar de minería de datos para múltiples industrias, del Inglés, Cross Industry Standard Process for Data Mining) se exploran distintos algoritmos utilizando telemetría real, obteniendo además información importante sobre las limitaciones en la persistencia y recuperación.

Desarrollo y resultados obtenidos

Previo al análisis de diferentes algoritmos se exploraron modernas técnicas de almacenamiento para series de tiempo que pretenden conjugar las mejores características de los RDBMS (Sistema de base de datos relacional del Inglés, Relational Database Management System) y las soluciones genéricamente conocidas como NoSql (No solo SQL, del Inglés, Not only SQL). De esta exploración se desprendieron decisiones de diseño respecto al UGS, pero además el trabajo cuestiona experimentalmente la conveniencia de trabajar con soluciones híbridas ofreciendo ventajas de limitado alcance en solo algunos escenarios.

Se exploraron además varios algoritmos para obtener fallas introducidas artificialmente sobre conjuntos de datos de telemetría real, obteniendo resultados alentadores en algunos casos. El uso de la mezcla gaussiana se destacó entre otros modelos ofreciendo buenos resultados en conjuntos de datos fuertemente correlacionados.

Las principales limitaciones del estudio radican en la ausencia de tener un conjunto de datos donde las fallas estén etiquetadas. Los fabricantes u operadores de estos sistemas son reacios a publicar información que exponga posibles fallas de diseño u operación y por tanto las anomalías deben ser introducidas artificialmente.

Conclusiones

Se desprende del trabajo (Soligo, Ierache, & Merkel, Telemetría de altas prestaciones sobre base de datos de serie de tiempos, 2020) los pocos argumentos que justifican el almacenamiento de telemetría satelital sobre un sistema híbrido. Se exploraron varios escenarios sin obtener mejoras relevantes y cuando las hubo, estas estaban circunscritas a un caso particular.

El uso de minería de datos y aprendizaje automático puede representar un salto en los sistemas de control de salud, aunque requiere de fuentes de datos estables y con alta correlación. Será necesario para avanzar en estas líneas tener conjuntos de datos etiquetados aun no disponibles. Los conjuntos de datos de telemetría satelital públicos donde se reconocen fallos son de una única variable o varias sin relación entre ellas, esto es la principal limitación para el avance en esta línea.

Publicaciones y/o transferencias realizadas

Escuela de Informática CACIC 2020

Curso Python de Cero a Django

Durante el curso, dictado por Pablo Soligo y German Merkel se utilizó material, datos y procesos derivados de los proyectos C211 y C230.

Reddi 2021

Artículo: Telemetría de altas prestaciones sobre base de datos de serie de tiempos.

WICC 2021

Artículo: Investigación, desarrollo y publicación de un prototipo de segmento terreno satelital.

CACIC 2021

Artículo: Detección de Anomalías en Segmento Terreno Satelital Aplicando Modelo de Mezcla Gaussiana y Rolling Means al Subsistema de Potencia.

Prototipos

Bibliografía

- Polson, S. (2019). *Unsupervised Machine Learning for Spacecraft Anomaly Detection in WebTCAD*. Delaware: University of Delaware, Laboratory for Atmospheric and Space Physics.
- scikit-learn Machine Learning in Python. (s.f.).
- Soligo, P., & Ierache, J. S. (2019). Arquitectura de segmento terreno satelital adaptada para el control de límites de telemetría dinámicos. *XXV Congreso Argentino de Ciencias de la Computación (La Plata, 2019)*.
- Soligo, P., Ierache, J. S., & Merkel, G. (2020). Telemetría de altas prestaciones sobre base de datos de serie de tiempos.
- Yairi, T., Kawahara, Y., Fujimaki, R., Sato, Y., & Machida, K. (2006). Telemetry-mining: a machine learning approach to anomaly detection and fault diagnosis for space systems. *Space Mission Challenges for Information Technology, 2006. SMC-IT 2006. Second IEEE International Conference on*, (págs. 8--pp).
- Yairi, T., Nakatsugawa, M., Hori, K., Nakasuka, S., Machida, K., & Ishihama, N. (2004). Adaptive limit checking for spacecraft telemetry data using regression tree learning. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, 6, págs. 5130-5135.
- Yairi, T., Takeishi, N., Oda, T., Nakajima, Y., Nishimura, N., & Takata, N. (2017). A Data-driven Health Monitoring Method for Satellite Housekeeping Data based on Probabilistic Clustering and Dimensionality Reduction. *IEEE Transactions on Aerospace and Electronic Systems*.