

Resumen de pasantías de investigación

Pasantías :: Instituto de Computación

Setiembre 2024

Formalización de sistemas críticos

(3 meses :: pasantía de investigación)

Referente: Carlos Luna cluna@fing.edu.uy, Grupo de Seguridad Informática

Muchos sistemas críticos pueden ser modelados y formalizados como máquinas de estados. Esta pasantía propone introducir a la o el estudiante en la especificación formal de un sistema de seguridad, analizando alguna propiedad relevante, usando el asistente de pruebas Coq.

Algoritmos de machine learning basados en bosques aleatorios para estimar densidades en problemas oceanográficos

(3 meses :: pasantía de investigación)

Referente: Héctor Cancela cancela@fing.edu.uy, Investigación Operativa

La estadística tradicional plantea la predicción de una variable de respuesta generalmente univariada. La mayoría de los algoritmos de aprendizaje automático también tienen como objetivo la predicción univariada. Sin embargo, es común predecir estructuras o respuestas más complejas y su abordaje no es trivial en términos analíticos y computacionales. En particular, se ha desarrollado un algoritmo basado en árboles de clasificación y regresión (CART) para predecir densidades (Nerini y Gattas 2007). Los CART son herramientas de fácil interpretación y muy útiles para la gestión pero inestables. En ese sentido, un algoritmo basado en un conjunto de árboles generados a partir de muestras bootstrap conocido como bosques aleatorios (Random Forest; RF, Breiman 2001) ha sido muy útil para mejorar la estabilidad y la performance de este tipo de modelos. Sin embargo, este tipo de algoritmo no ha sido aplicado para predecir densidades. La presente propuesta propone generar un algoritmo basado en RF para predecir densidades (RFd). En esta pasantía se deberá desarrollar un algoritmo eficiente para implementar los RFd y evaluar su performance en datos simulados y aplicarlo a datos oceanográficos reales. Se deberá generar una implementación del mismo en el software libre R

(<https://cran.r-project.org/>). La dirección del trabajo será realizada en forma conjunta por Héctor Cancela (INCO) y Angel Segura (CURE).

Aprendizaje de un sistema binario estocástico

(3 meses :: pasantía de investigación)

Referentes: Pablo Romero y Héctor Cancela promero@fing.edu.uy, Investigación Operativa

El desarrollo de la presente pasantía de investigación requiere conocimientos de Programación, Lógica e Introducción a la Investigación de Operaciones.

Un sistema binario estocástico es un sistema que posee múltiples componentes que pueden fallar en forma independiente y con idéntica probabilidad. La confiabilidad de un sistema binario estocástico es la probabilidad de que el sistema funcione correctamente. El estudio de sistemas binarios estocásticos presenta potenciales aplicaciones en sistemas de comunicaciones y en la asistencia a la toma de decisiones.

El estudio de SBS presenta desafíos matemáticos y computacionales. El primer desafío es que el cálculo de la confiabilidad de un SBS es en general un problema NP-Difícil. El segundo es que la memoria necesaria para almacenar a la estructura de un SBS en forma de tabla es exponencial en la cantidad de componentes, por lo que no es plausible la representación de una estructura en forma de tabla para sistemas con una cantidad de componentes muy numerosa.

El objetivo de la presente pasantía es familiarizarse con el concepto de sistema binario estocástico. Analizaremos sistemas que se representan como un grafo, donde el objetivo consiste en comunicar 2 o más vértices especiales y los restantes vértices pueden fallar. Por último, estudiaremos el problema del aprendizaje de un sistema desconocido.

Tiempo de difusión en redes

(3 meses :: pasantía de investigación)

Referentes: Pablo Romero y Héctor Cancela promero@fing.edu.uy, Investigación Operativa

El estudio del tiempo mínimo de difusión en redes presenta interés en el diseño de sistemas distribuidos y en telecomunicaciones. En Facultad de Ingeniería se ha desarrollado un Proyecto de grado relativo a este tema,

y recientemente en Montreal se ha defendido una tesis doctoral que presente progresos en el área. Se recomienda que quien elija esta propuesta posea conocimientos en programación, teoría de grafos e investigación operativa.

Un nodo de la red posee un mensaje que desea comunicar a todos los nodos de la red. En cada unidad de tiempo, cada nodo que posee el mensaje puede reenviar el mensaje a un nodo vecino. El tiempo mínimo de difusión iniciando de un nodo es el menor tiempo que permite alcanzar a todos los nodos de la red. El objetivo de esta pasantía es determinar el tiempo mínimo de difusión de redes especiales.

Desarrollo de un modelo de Redes Neuronales de Grafo para la predicción de latencia en aplicaciones basadas en microservicios

(3 meses :: pasantía de investigación)

Referentes: Matias Richart y Javier Baliosian mrichart@fing.edu.uy, Network Management / Artificial Intelligence

La computación en la nube y la computación de borde han revolucionado la forma en que se manejan las tareas informáticas, ofreciendo flexibilidad y eficiencia. A la vez, la escala y complejidad de esas tareas ha motivado el uso de la arquitectura de microservicios que permite desarrollar sistemas complejos de forma modular, reutilizable y escalable. Los microservicios permiten también una asignación más fina de los recursos disponibles, pero añaden complejidad a la gestión.

Para contribuir al uso eficiente de los recursos, es deseable que las aplicaciones sean elásticas. Es decir, que tengan la capacidad de adaptar los recursos asignados a un sistema de forma dinámica, autónoma y óptima a la carga de trabajo variable con el objetivo de minimizar la cantidad de recursos necesarios mientras se cumplen los requisitos de calidad de servicio. Sin embargo, esta combinación de características se ha vuelto extremadamente compleja de gestionar. Atacar el problema de la elasticidad en aplicaciones basadas en microservicios implica una serie de retos que no se presentan cuando se trabaja con aplicaciones monolíticas; en lugar de asignar recursos a un servicio singular, el controlador de la elasticidad debe evaluar ahora las dependencias entre cualquier par de microservicios para garantizar la calidad del servicio.

Actualmente estamos desarrollando un proyecto de investigación que aborda el problema de la elasticidad de aplicaciones basadas en microservicios y desplegadas en el continuo de la nube, entre los centros de datos y el borde de la red. Parte de la solución de este problema implica contar con un predictor del rendimiento de una aplicación de microservicios de forma de poder tomar decisiones en tiempo real. Para el desarrollo de este predictor proponemos un enfoque

novedoso que hace uso de Redes Neuronales de Grafos (GNN) en combinación con Redes de Colas por Niveles (LQN).

Actualmente se cuenta con un diseño del modelo y un prototipo implementado en un framework de GNN. El objetivo de esta pasantía es incorporarse a este proyecto e implementar el modelo de GNN ya diseñado utilizando el lenguaje Python y la biblioteca TensorFlow. Como resultados finales del trabajo, se espera contar con una implementación del modelo y su correspondiente documentación, así como realizar el entrenamiento y su posterior evaluación. Al final de la pasantía se espera que el/la pasante haya experimentado de cerca el trabajo en un proyecto de investigación y adquirido conocimientos en el trabajo con algoritmos de aprendizaje automático.

Calibración de parámetros en metaheurísticas

(3 meses :: pasantía de investigación)

Referentes: Martín Pedemonte mpedemon@fing.edu.uy, Computación de alto desempeño aplicado

Las metaheurísticas son una familia de algoritmos aproximados para la resolución de problemas de optimización y búsqueda de propósito general. Son capaces de encontrar soluciones de alta calidad (muchas veces óptimas) a problemas de optimización en tiempos de ejecución razonables. A diferencia de las técnicas heurísticas que implican diseñar un procedimiento específico para cada problema de optimización, las metaheurísticas son esquemas generales de algoritmos de optimización que pueden utilizarse para abordar una amplia gama de problemas. En general, las técnicas metaheurísticas pueden aplicarse a las particularidades de cada problema con relativamente pocas modificaciones.

Las metaheurísticas exploran el espacio de búsqueda de manera eficiente, con el objetivo de encontrar soluciones factibles de alta calidad. El diseño de un algoritmo de este tipo debe tener en cuenta un equilibrio adecuado entre la exploración del espacio de búsqueda y la explotación de las mejores soluciones encontradas. También debe tener la capacidad de no quedarse estancada en regiones particulares del espacio de búsqueda, especialmente evitando quedar atrapada en óptimos locales. Algunas de las técnicas más estudiadas son los Algoritmos Evolutivos y los algoritmos basados en inteligencia de enjambre, como Ant Colony Optimization.

La utilización de metaheurísticas para la resolución de un problema concreto involucra una etapa previa de ajuste o calibración de parámetros de la propia técnica que se va a usar. En general, el ajuste de parámetros o calibración se suele realizar siguiendo un enfoque de fuerza bruta y/o con poco rigor estadístico. En los últimos años, han surgido herramientas para la configuración automática de algoritmos, de forma de encontrar automáticamente los mejores ajustes de parámetros de un método. En particular, irace es un paquete de software que implementa una serie de procedimientos de configuración automática que buscan dotar a la etapa de calibración de parámetros de una mayor rigurosidad. En contraposición, el uso de

este tipo de herramientas puede aumentar el esfuerzo computacional dedicado a dicha etapa, por lo que también es importante comprender la relación existente entre el tiempo de ejecución requerido durante la etapa de calibración de parámetros de los algoritmos y la calidad de las soluciones que se obtienen.

El objetivo de la presente pasantía de investigación es evaluar el funcionamiento de la herramienta de configuración automática irace, así como evaluar la relación entre la calidad de las soluciones obtenidas por una metaheurística con relación al tiempo de ejecución destinado a la etapa de calibración de parámetros.

Aplicación de un enfoque neuro simbólico basado en grafos para enriquecer ontologías.

(3 meses :: pasantía de investigación)

Referentes: Edelweis Rohrer erohrer@fing.edu.uy, Sistemas de Información Semántica

Las ontologías son los enfoques formales de modelado por excelencia, por su capacidad para conceptualizar diferentes dominios, asegurar la consistencia del modelo e inferir nuevo conocimiento, a través de la aplicación de razonamiento deductivo. Las redes de ontologías, que integran ontologías de diferentes dominios, son fundamentales porque favorecen el intercambio de datos entre diferentes sistemas, facilitando la interoperabilidad semántica. Los enfoques neuro simbólicos integran modelos de aprendizaje automático con ontologías, aplicando mecanismos de razonamiento híbridos inductivo-deductivo. En el área de ingeniería de ontologías, estos enfoques están resultando de gran utilidad, ya que enriquecen las ontologías al predecir nuevas relaciones y clasificar instancias. Al mismo tiempo, los modelos de aprendizaje automático se benefician de las ontologías, que les proveen de conocimiento para comprender mejor el contexto y las conexiones entre términos, mejorando su capacidad de generar conocimiento relevante.

Algunos enfoques neuro simbólicos “basados en grafos”, aplican una estrategia en dos pasos: ontología- grafo, que transforma una ontología a una estructura de grafo, y grafo-espacio vectorial, que embebe el grafo resultante en el espacio vectorial del modelo de aprendizaje automático. El conocimiento así generado es incorporado a la ontología recorriendo el camino inverso: espacio vectorial-grafo y grafo-ontología.

En el marco de un estudio del uso de diferentes enfoques neuro simbólicos en el área de ingeniería de ontologías, el objetivo de esta pasantía es la aplicación de un enfoque neuro simbólico basado en grafos a seleccionar al desarrollo de una red de ontologías, para evaluar los resultados obtenidos en instancias posteriores, comparándolos con otros enfoques neuro simbólicos y con la aplicación del enfoque tradicional de desarrollo manual.

Los resultados esperados son:

- Aplicación del enfoque neuro simbólico basado en grafos a una versión inicial de una red de ontologías, para enriquecer las ontologías y predecir nuevas relaciones entre

ellas, asegurando que la red obtenida preserva la consistencia lógica y semántica (evaluada por un experto del dominio).

- Registro de los resultados obtenidos y de las actividades desarrolladas.

ADA-Partituras

(3 meses :: pasantía de investigación)

Referentes: Regina Motz rmotz@fing.edu.uy, Sistemas de Información Semántica

Esta pasantía propone incorporarse a un equipo interdisciplinario que se encuentra desarrollando la aplicación ADA: App de Dictados Armónicos.

Esta aplicación móvil está siendo desarrollada siguiendo una metodología de co-diseño con el grupo CIEDA (Construcción Interdisciplinaria de la Educación Digital Abierta) del Espacio Interdisciplinario del cual forma parte el grupo SIS del InCo, profesionales de Ciencias de la Educación y docentes del Departamento de Teoría y Composición del Instituto de Música de la Facultad de Artes.

La aplicación es un generador de dictados armónicos que ofrece dictados a los estudiantes para realizar entrenamiento auditivo de acuerdo a pautas establecidas por los docentes en relación a los elementos importantes para cada curso específico. El estudiante puede entrenarse sobre los mismos elementos musicales con el mismo grado de dificultad a través de ilimitados dictados diferentes que genera la aplicación.

La primera versión de ADA está disponible en las tiendas de Google y de Apple Store y puedes encontrar su presentación en <https://ada-web-two.vercel.app/>

El módulo en el que se ofrece trabajar durante esta pasantía es en la corrección automática de la transcripción de los dictados. Esto permitirá proporcionar una retroalimentación inmediata y detallada sobre los errores en la escritura de las partituras a partir del dictado musical. La aplicación de algoritmos avanzados para analizar el rendimiento del estudiante en tiempo real permitirá a los usuarios establecer metas personales, monitorear su progreso y ajustar su estrategia de estudio en función de los datos obtenidos.

Para la realización de esta pasantía es requisito tener disponibles dos horas semanales de coordinación con todo el equipo interdisciplinario que se realizan por Zoom en el horario central (se coordina día y hora pero es entre las 8:00hrs y las 17:00hrs). El resto del horario de dedicación es flexible.

Desarrollar una plataforma de emulación de redes ópticas

(3 meses :: pasantía de investigación)

Referentes: Alberto Castro acastro@fing.edu.uy, Network Management / Artificial Intelligence

Las comunicaciones juegan un rol muy importante en las interacciones humanas, en la economía, en la educación, en el acceso y democratización de distintos servicios e información. Con una amplia infraestructura física desplegada, las redes ópticas han desempeñado un papel fundamental en las comunicaciones de banda ancha de extremo a extremo y son consideradas la base para las futuras comunicaciones en la sociedad moderna. Se espera que las redes ópticas de futura generación brinden servicios de banda ancha con capacidad masiva, latencias más bajas y confiabilidad mejorada soportando varias aplicaciones de ultra alto ancho de banda: cloud/edge networking, servicios de video 8K y las incipientes comunicaciones 3D (hologramas) que vienen de la mano de la digitalización/replicación del mundo que nos rodea (gemelos digitales). Las redes ópticas vienen teniendo un rol cada vez más importante en las comunicaciones móviles y serán clave para las redes 5G/6G dados sus requerimientos de ancho de banda, cobertura, sincronización y bajas latencias. El éxito de las futuras comunicaciones móviles no dependerá únicamente de las nuevas tecnologías de acceso por radio, sino que dependerá fuertemente de las redes ópticas y de la capacidad de cómputo de los extremos.

El objetivo de esta pasantía es desarrollar una plataforma de emulación de redes ópticas compatible con la arquitectura OpenROADM.

OpenROADM (Open Reconfigurable Optical Add/Drop Multiplexer) es un estándar abierto que está transformando la industria de las redes ópticas. Permite la interoperabilidad entre equipos de diferentes fabricantes, reduciendo costos y aumentando la flexibilidad en el diseño de redes. Se propone desarrollar un emulador que:

1. Implemente ROADMs (Reconfigurable Optical Add/Drop Multiplexers) utilizando el simulador de dispositivos Honeynode, el cual es compatible con OpenROADM.
2. Funcione en contenedores Docker para facilitar el despliegue y la escalabilidad.
3. Sea compatible con plataformas como Kathará o Containerlab para realizar experimentos a gran escala (ambas plataformas utilizan Docker).
4. Proporcione un entorno de pruebas en tiempo real para nuevas arquitecturas de gestión y control de redes ópticas, algoritmos de optimización y algoritmos de aprendizaje automático.

La plataforma permitirá a investigadores, estudiantes y profesionales de la industria experimentar con redes ópticas complejas sin necesidad de hardware costoso. El emulador también será una herramienta educativa eficaz para la enseñanza hands-on de redes ópticas.

Implementación de un compilador MateFun usando AspectAG

(3 meses :: pasantía de investigación)

Referentes: Marcos Viera, mviera@fing.edu.uy, Grupo de Métodos Formales

Las Gramáticas de Atributos (AG) son un formalismo que permite describir la semántica de un lenguaje de programación de manera modular, decorando las producciones de una gramática libre de contexto con valores (atributos). AspectAG es una biblioteca de Haskell, desarrollada por el grupo de Programación Funcional del InCo, que implementa un lenguaje de AG fuertemente tipado. Por su parte, MateFun es un lenguaje de programación funcional, también desarrollado en el InCo, dirigido al aprendizaje de funciones matemáticas por adolescentes. El objetivo de esta pasantía es implementar un compilador para parte esencial de MateFun usando AspectAG.