

Procesamento aproximado

O incremento no volume, velocidade e variedade das fontes de datos que teñen que ser analizadas en procesos de toma de decisións está alcanzando límites insospeitados. Unha grande parte destes datos son xerados automaticamente por dispositivos sensores (na industria, no sector saúde, na monitorización ambiental, no crowdsensing, nas cidades intelixentes, etc.). As tecnoloxías Big Data permitiron a xeración de novo coñecemento a partir desta inmensa cantidade de datos a costa de incrementar os recursos computacionais. A pesar dos esforzos intensivos de investigación nesta área, entre os que se inclúen investigación en explotación de datos, paralelismo, indexación, comprensión de datos, procesamento eficiente de consultas e optimización etc., os tempos de resposta que requiren as tarefas de explotación de datos son pouco satisfactorios en moitos casos.

O procesamento aproximado ten como obxectivo proporcionar resultados aproximados ás necesidades dos usuarios nunha fracción do tempo requirido polos métodos de procesamento tradicional, permitindo a exploración interactiva eficiente de conxuntos de datos moi grandes mediante o uso de recursos computacionais limitados. Este ambicioso obxectivo alcánzase por medio da explotación efectiva de resumos dos datos, a partir de histogramas, bosquexos e mostraxes.

Obxectivos científicos

A partir do anterior, as liñas principais deste programa de investigación son as seguintes:

- Investigar en técnicas de procesamento aproximado para 'lagos de datos', i.e., repositorios de datos de grande tamaño e heteroxeneidade almacenados na súa forma natural (relacionais, csv, logs, xml, json, documentos, emails, imaxes, audio, etc.).
- Procesamento en tempo real de streams de datos.
- Tecnoloxías de aprendizaxe sobre bases de datos que permiten a construción interactiva dun modelo do dataset utilizando información extraída de procesos previos posiblemente a partir de técnicas de machine learning, as cales poden ser empregadas para incrementar a velocidade do novo procesado.
- Visualización Optimista, que produce resultados aproximados nun curto lapso de tempo, o que permite aos analistas de datos a realización de observacións sobre os valores aproximados, que poden ser contrastados posteriormente contra resultados máis precisos.

Obxectivos tecnolóxicos

Aplicación de procesamento aproximado para a exploración interactiva de data sets de información de sensores de grande tamaño nos ámbitos de:

- Xestión de información xeoespacial e medioambiental
- Optimización de procesos industriais
- Edificios intelixentes e cidades intelixentes

Queres saber máis?

Para colaboracións, visitas, etc. contacta con:

José Ramón Ríos Viqueira

jrr.viqueira@usc.es

[@RiosViqueira](https://twitter.com/RiosViqueira)

[jrviqueira](https://www.linkedin.com/in/jrviqueira)

INVESTIGADORES ASOCIADOS

José Manuel Cotos Yáñez
Julián Flores González
David Mera Pérez
José Ramón Ríos Viqueira
José Ángel Taboada González

PROXECTOS DE INVESTIGACIÓN

FIBRATIC: Sistemas Intelixentes para a Industria 4.0

MiniStor: MiniStor-MINimal Size Thermal and Electrical Energy STORage System for In-Situ Residential Installation

MAGIST-ELA: Xeoprocesamento a gran escala para análise exploratorio e baseado en aprendizaxe

[Ver todos](#)

PUBLICACIONES

Efficient access methods for very large distributed graph databases
Information Sciences, 2021

Perspectiva de género en Inteligencia Artificial
I Congreso Internacional Tecnologías I+D+i de igualdad: soluciones, perspectivas y retos, 2021

Exploratory study of the impact of project domain and size category on the detection of the God class design smell
Software Quality Journal, 2021

[Ver todas](#)