

PRI06A227 - Optimización de sistemas de clasificación mediante GA y arquitectura BOINC

Título: Optimización de sistemas de clasificación mediante Algoritmos Genéticos y arquitectura BOINC

Entidad financiadora:

Junta de Extremadura

Duración: Enero 2007 - Diciembre 2008 Investigador principal: Carlos J. García-Orellana

Participantes CAPI: 5

Descripción: Con el presente proyecto tenemos como objetivo evaluar el uso de la infraestructura informática y telemática existente en Extremadura a nivel institucional (en colegios, hospitales, universidad, administración) para resolver problemas científicos de fuerte carga computacional. La propuesta pretende estudiar la viabilidad y plantear soluciones a los problemas de este tipo de computación en procesos complejos de optimización.

Para ello, planteamos el desarrollo de herramientas de SoftComputing para su utilización en entornos del tipo PRC (Public Resource Computing). El proyecto se basa en el uso de la arquitectura BOINC (Berkeley Open Infrastructure for Network Computing), ya que ésta es la arquitectura que estará presente en el entorno Extreme.

En particular proponemos la adecuación y desarrollo de herramientas para abordar la optimización de sistemas de clasificación mediante Algoritmos Genéticos en problemas, generalmente, de análisis de imágenes. La optimización de sistemas de clasificación es un problema complejo, con numerosos parámetros a ajustar (en la extracción y selección de características y en el clasificador en sí) y con enorme coste computacional y relativamente baja carga de comunicación. Este alto ratio computación/comunicación lo hacen adecuado para la arquitectura BOINC.

Para desarrollar el proyecto proponemos tres objetivos secuenciales: primero comprobaremos el funcionamiento de la simulación de clasificadores neuronales bajo BOINC (ya tenemos migrada la aplicación), segundo abordaremos la ejecución de Algoritmos Genéticos (GA) para optimizar problemas con funciones de evaluación costosas desde el punto de vista computacional. Como tercer objetivo nos planteamos la inclusión de un modelo basado en el concepto de 'Calidad de Servicio'

(QoS), en el proceso de simulación, mediante el uso de recursos dedicados y controlados, es decir, mediante el uso, junto con BOINC, de un pequeño cluster de apoyo.

Las aplicaciones del proyecto son numerosas. Nosotros nos centraremos durante su ejecución en aplicarlo a dos de los problemas con los que trabaja nuestro grupo de investigación: la clasificación de cubierta nubosa y la clasificación de clusters de microcalcificaciones en mamografías. Si los resultados del proyecto son satisfactorios, el sistema desarrollado será clave para el desarrollo de nuevos problemas complejos de clasificación. Pero eso será solamente el principio, las herramientas

se diseñarán y desarrollarán pensando en su aplicación a otros problemas de optimización y se harán públicas. Esto implica que otros grupos de trabajo podrán utilizar la base desarrollada en sus propios problemas, aprovechando la experiencia adquirida por nuestro grupo.

Este último aspecto, el que sirva de base para los demás, es muy importante en el planteamiento del proyecto, ya que mantener una infraestructura de este tipo solamente tiene sentido si existen usuarios que lo utilicen, bastantes usuarios. Por ello, pretendemos dotar de unos contenidos y aplicaciones iniciales a la infraestructura que la Junta de Extremadura está poniendo en marcha, y demostrar a los demás la potencialidad de esta tecnología de vanguardia basada en BOINC.