



Confirmation d'amas de galaxies par techniques d'apprentissage

Spécialité Astrophysique

Niveau d'étude Bac+5

Formation Master 2

Unité d'accueil

Candidature avant le 01/05/2017

Durée 3 mois

Poursuite possible en thèse non

Contact [Tarrío Alonso Paula](#)
+33 1 69 08 37 87
paula.tarrio-alonso@cea.fr

Résumé

Le but du projet est de tester et d'ajuster des algorithmes d'apprentissage pour automatiser la présélection de candidats amas de galaxies en utilisant des données photométriques optiques. En particulier, l'étudiant utilisera l'information photométrique de WISE et de SDSS autour d'amas de galaxies connus et confirmés à différents redshifts pour construire un classificateur qui sera en mesure de séparer la population d'amas en différentes classes (principalement, haut et bas redshift).

Sujet détaillé

Les amas de galaxies sont des objets extrêmement intéressants d'un point de vue cosmologique qui nous aident à comprendre comment l'Univers évolue. En particulier, leur abondance au cours du temps cosmique est étroitement liée au contenu en matière et énergie de l'Univers et à son évolution. Dans ce contexte, la découverte d'amas massifs à haut redshift est très utile.

Les amas de galaxies peuvent être détectés en utilisant des données d'observation à différentes longueurs d'onde, depuis les micro-ondes (avec par exemple, Planck) aux rayons X (avec par exemple, ROSAT). De manière générale, les techniques de détection d'amas produisent une grande liste de candidats, dont certains peuvent ne pas être de véritables amas. Puisque faire un suivi de tous ces candidats (i.e. confirmer qu'ils sont de véritables amas et extraire certaines de leurs caractéristiques, telles que la masse, la température, etc.) serait une tâche très coûteuse en terme de temps de télescope, une présélection ou classement des candidats les plus prometteurs est généralement effectuée.

L'étudiant sera impliqué dans un projet en cours pour détecter des amas de galaxies à haut redshift en combinant les observations Planck et ROSAT. En particulier, le projet a produit une liste de candidats d'amas et un de nos objectifs est de présélectionner les candidats qui sont le plus probablement des amas à haut redshift. Afin de faire cette présélection, les données optiques de SDSS et WISE peuvent être utilisées. Par exemple, les galaxies d'amas massifs à $z > 0,7$ devraient être brillantes dans le canal de 3.4 μ m de WISE, et faibles ou invisibles dans la bande i de SDSS. En recherchant des surdensités de sources avec ces critères, on peut identifier des candidats d'amas avec un haut niveau de confiance, et on peut obtenir une première estimation du redshift de l'amas à partir de la photométrie. L'étudiant testera et ajustera des algorithmes d'apprentissage pour automatiser cette présélection. En particulier, il

utilisera les informations photométriques de WISE et SDSS autour d'amas de galaxies connus et confirmés à différents redshifts pour construire un classificateur qui sera en mesure de séparer la population d'amas en différentes classes (bas z et haut z , avec une possible extension à différents plages de redshift, et / ou à différents plages de masse).

Le projet est co-encadré par Jean-Baptiste Melin (jean-baptiste.melin@cea.fr)

Mots clés

amas des galaxies, traitement de données, techniques d'apprentissage

Compétences

Maîtrise d'un langage de programmation

Logiciels

MAC OS and/or GNU-LINUX, IDL or PYTHON or MATLAB

Confirmation of galaxy clusters using machine learning

Summary

The goal of the project is to test and adjust machine learning algorithms to automate the pre-selection of galaxy cluster candidates using optical photometric data. In particular, the student will use WISE and SDSS photometric information around known and confirmed galaxy clusters at different redshifts to build a classifier that will be able to separate the cluster population into different classes (mainly, high-redshift and low-redshift).

Full description

Galaxy clusters are very interesting objects that can help us understanding how the universe evolves, since their abundance over cosmic time is tightly related to the universe content and its evolution. In this context, discovering high-redshift and massive clusters is extremely useful.

Galaxy clusters can be detected using observational data sets at different wavelengths, from microwaves (e.g. Planck) to X-rays (e.g. ROSAT). Generally speaking, cluster detection techniques may produce a large list of cluster candidates, some of which may not be real clusters. Since doing a follow-up of all the candidates (to confirm if they are real clusters and to extract some of their characteristics, such as mass, temperature, etc.) would be a very expensive task in terms of telescope time, a pre-selection or grading of the most promising candidates is usually performed.

The student will be involved in a current project to detect high-redshift galaxy clusters by combining observations from Planck and ROSAT surveys. In particular, the project has produced a list of cluster candidates and one of our goals is to pre-select the candidates that are most-likely high-redshift clusters. In order to do this pre-selection, optical data from the Sloan Digital Sky Survey (SDSS) and NASAs WISE survey can be used. For example, massive cluster galaxies at $z > 0.7$ are expected to be bright in the 3.4 μ m WISE channel, and faint or invisible in the SDSS i-band filter. By searching for over-densities of sources with these criteria, high-confidence cluster candidates can be identified, and a first estimate of the cluster redshift can be obtained from the photometry.

The student will test and adjust machine learning algorithms to automate this pre-selection. In particular, he will use WISE and SDSS photometric information around known and confirmed galaxy clusters at different redshifts to build a classifier that will be able to separate the cluster population into different classes (low- z and high- z , with a possible extension to different redshift bins, and/or to different mass bins).

The project is co-supervised by Jean-Baptiste Melin (jean-baptiste.melin@cea.fr)

Keywords

galaxy clusters, data processing, machine learning

Skills

Knowledge of a programming language

Softwares

MAC OS and/or GNU-LINUX, IDL or PYTHON or MATLAB