

W19b 全天X線監視装置 MAXI の観測データの再解析

西尾颯斗, 根来均 (日本大学)

全天X線監視装置 MAXI により、これまで多くの突発現象や突発天体が発見されてきた。しかし、運用初期のデータなど一部のデータに対してはこれまで十分な突発天体探査が行われておらず、突発現象を見逃してきた可能性がある。そこで我々は、過去 15 年以上にわたる全データを再解析するため、これまでのデータをデータベース (PostgreSQL) に登録する作業を行っている。前回の年会では、元データの一部で確認された重複データを処理する際に、データベースに一括登録するための CSV ファイルの生成に時間がかかることから、検出された重複データを Python ライブラリ Pandas の DataFrame から 1 行ずつ削除するのではなく、データベースに登録された重複データを削除することにより、重複データの処理が 1530 倍高速化されたことなどを発表した。

今回、DataFrame の高速処理が可能な Python ライブラリの Polars (<https://pola.rs>) を導入することで、DataFrame を CSV ファイルに書き出す際に用いられていた Pandas の `to_csv` メソッドの実行時間に比べて、書き出し時間が約 25 倍速くなり、ようやくハードウェアから期待される限界の処理速度に近くなった。また、一部の SSC のデータには 2 つの元のファイルに跨る 10 万レコード以上の重複したデータがあることがわかった。それらをデータベースから削除することにより大量の未使用領域が生じるため、定期的に VACUUM 処理を行いデータベースの肥大化を抑制した。現在、再解析に向けて解析システムへの機械学習の導入に取り組んでいる。機械学習には日大の朝倉敬介の手法を基に、教師あり学習の回帰分析を用いる。これまで試験的に導入している、トリガー画像のフィッティング解析システムの改良に加え、そのフィッティングパラメータを用いて真偽判定を行う予定である。本講演では、これらのデータ登録システムの改良と解析システムの現状について発表する。