

分散配列データベースを用いた月地震データの比較分析ツールの開発 Development of Analysis Tool for Moonquake Data Using Distributed Array Database

酒井 健[†] 木谷友哉[†] 山本幸生[‡] 山田竜平[§] 横山昌平[¶]
Takeru Sakai Tomoya Kitani Yukio Yamamoto Ryuhei Yamada Shohei Yokoyama

1. 概要

1969年から1977年の約7年半の間、NASAのアポロミッションにおいて、月面に設置された地震計を用いた連続観測により、膨大な量の月地震データが得られた。このデータは月地震に関する唯一のデータとして様々な解析が行われている。本研究では、時系列データの扱いに適している配列型データベースであるSciDBをバックエンドに用い、月地震データを比較分析するツールの開発を行った。このツールはユーザが選択した2つの期間の地震波形データを可視化する他、波形間距離やヒストグラムなど分析に必要な情報をオンデマンドに表示する。本システムは月を研究する専門家のニーズに合わせて開発した。

2. はじめに

NASAのアポロミッションによってALSEP(Apollo Lunar Surface Experiment Package)という観測装置群が設置されている。[1]そのALSEPの中に含まれている地震計を用いた月面上で月地震(以後、月震と表記する)の連続観測を行うPSE(Passive Seismic Experiment)と呼ばれる実験が行われた。その実験により約7年半(1969~1977)にわたる月震記録が得られ、月における地震が初めて観測された。

現在その波形データはすべてWebにて公開され、データの閲覧、取得が可能となっている。[2]* PSEによって得られたデータは、アポロミッションから40年経った現在も解析が行われており、月震の発生原因、活動度、月の内部構造などの研究において使用され、現在も多くの知見が得られている。

本研究では月震データ解析を行う専門家の補助することを目的として、月震可視化システムを構築する。

また、月震データは関係データベースシステム(RDBMS)で扱うには不適切である。一般にRDBMSは順序のないデータのみを扱う。しかし、月震データのようなセンサデータは時系列により順序が存在するため、RDBMSで順序を表すためには時系列にソートする必要がある。このようにRDBMSでは時系列データを扱うにはコストがかかってしまう。近年では、順序付けデータを専門に扱うデータベースシステムが多く開発されている。例えば、SciDB(Science-oriented DBMS)がある。SciDBはParadigm社**によって開発されたアレイデータベースシステムであり、センサデータの

他に地理空間データや遺伝子配列データに使われている。また、分散処理や並列処理を行うアーキテクチャによってスケールアウトすることにより、多くのデータを格納し、アクセスできるという特徴がある。そのため、時系列データの扱いに適した配列型データベースSciDBを用いて時系列データの取り出しを高速化する。

3. 関連研究

月震の分類に関する研究として、Nakamura[3]の研究が挙げられる。月震波形に対して、相互相関係数を用いた最短距離法による階層的クラスタリングを行い、その結果に基づいて人手による月震分類を行っている。この研究による分類が現在の月震研究の基準となっている。

その後、様々な研究者によりNakamuraらの研究によって発見できなかった月震が発見されている。

Bulowら[4]の研究では、波形の前処理を改良することで、新たな震源の月震を多数発見している。

後藤ら[5]によるSOMを用いた月震の可視化システムでは、従来の月震波形の相互相関係数ではなく周波数成分を特徴量として教師なし学習のひとつであるSOMによる深発月震の分類を可視化し、月震解析基盤には分散処理などを用いてRDBMSの高速化を図った。また、この研究により、これまで人手によって行われていた分類は、機械学習を用いた分類によって、さらに容易になり、震源の分類基準の改良や新たな月震の発見を可能にした。

これらの研究から、従来の分類基準を改良することや、波形の前処理方法を改良することで、アポロミッションによるPSEで得られたデータには未だに発見されていない月震を発見できる可能性がある。そのためデータ解析する環境を整えることは必要であるといえる。また、SciDBとRDBMSのベンチマークを比較したPhilippeら[6]による研究では、SS-DBという新しいベンチマークテストを提案し、99GBと990GBの膨大な配列データを用いて、RDBMSであるMySQLとSciDBの性能を比較し、SciDBが配列データに適していることを示した。

これらの関連研究から、本研究ではSciDBを用いた月震分類研究のためのシステムを提案する。

4. 提案システム

4.1. 概要

本研究では専門家を手助けするための月震可視化システムを提案する。本システムは、SciDBを用いることでデータの取得をより容易にし、研究者が解析しやすい環境を構築することなど、月震研究の補助を目的とする。また、月震データは500GBを超える大規模

[†]静岡大学院

[‡]宇宙航空研究開発機構宇宙科学研究所, JAXA

[§]会津大学

[¶]首都大学東京

*<https://www.darts.isas.jaxa.jp/planet/seismology/apollo/app/>

**<http://www.paradigm4.com/>

な時系列データであり、13058件の月震を含んでいる。本システムはWebSocketを用いてサーバ側とクライアント側でデータの送受信を行う。サーバ側で月震データをクライアント側へ送信し、クライアント側で可視化を行う仕組みとなっている。クライアント側はWebブラウザへグラフを描画することで、研究者の手にデータがなくとも、容易にデータにアクセス、閲覧が可能になっている。必要であればデータをダウンロードすることも可能であり、可視化されたデータを見ながらさらに手で解析を進めることが可能である。

4.2. システム構成

システム構成は図1のようにになっている。データベースにSciDB、サーバ側はPython、クライアント側はJavaScriptで構築されている。ブラウザのフォームか

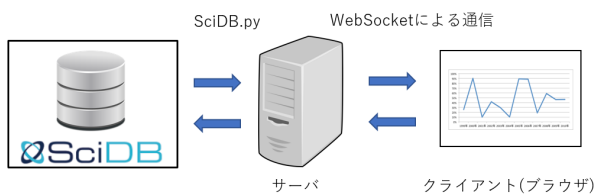


図1: システム構成図

ら取得したいデータの時刻を送信することで、サーバ側でその時刻分のデータをSciDBから取り出し、再度ブラウザへ送り返し、グラフを描画する仕組みとなっている。データの送受信はWebSocketを利用している。ブラウザへアクセスした際にクライアントとサーバが接続される。

4.3. 可視化

月震データは膨大であり、そのすべてを一度に描画するのは現実的ではないため、データの間引きを行い、サーバ側とクライアント側の送受信データ量を削減している。例えば、50,000セグメントのデータを50セグメント毎に代表値を1つ選択し、1,000セグメントへ削減している。ただし、この場合、月震データから得られる特徴量が同じ50セグメント内に複数存在しており、代表値として選択されなかった場合、残りのデータは破棄されるため、特徴量を見逃してしまうという課題がある。そこで、Highchart.jsのRangeSeriesという描画グラフを用いることで解決する。RangeSeriesは図2のようなグラフであり、2点間を塗りつぶして表示できるため、50セグメント毎の最大値と最小値を代表値として選択することで塗りつぶされた2点の範囲に値が収まるようになっている。複数グラフを同時に表示できるため、複数のグラフを比較することも可能であり、単一グラフの表示も可能である。その場合、グラフの凡例をクリックすることで切り替えることができる。また、マウスオーバーによってデータの値が数値で確認できる。

このようにして、送受信データ量を削減しつつ、より近似したグラフを描画する。その他にも周波数のヒストグラムや箱髭図なども描画可能であり、どのようなグラフが必要かはユーザが選択することができる。

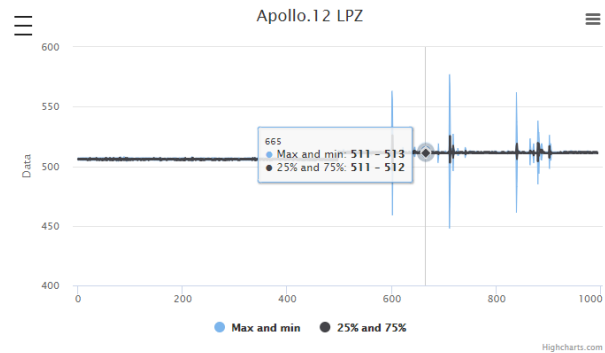


図2: RangeSeries グラフ

5. まとめと今後の課題

本研究では、専門家の月震分析を補助するための月震可視化システムを構築した。システムのバックエンドには時系列データを専門に扱う配列型データベースであるSciDBを用いた。今後は、専門家に本システムを実際に使用してもらい、フィードバックを得て、より専門家のニーズにあったシステムへと改良し続けたいと考えている。

参考文献

- [1] N. Toksöz, A. Dainty, C. Solomon, and K. Anderson. Structure of the moon. *Reviews of Geophysics and Space Physics*, Vol. 12, No. 4, pp. 539-565, 1974.
- [2] 山田竜平, 山本幸生, 桑村潤, and 中村吉雄. アポロ月地震データ公開システムの開発. *宇宙科学情報解析論文誌宇宙航空研究開発機構研究開発報告*, 1 (2012), 121-131
- [3] Nakamura, Y. New identification of deep moonquakes in the apollo lunar seismic data. *Physics of the Earth and Planetary Interiors* 139, 3 (2003), 197-205.
- [4] R. C. Bulow, C. L. Johnson, and PM Shearer. New events discovered in the apollo lunar seismic data. *Journal of Geophysical Research: Planets* (1991-2012), Vol. 110, No. E10, 2005.
- [5] 後藤康路, 山田竜平, 山本幸生, 横山昌平, and 石川博. 階層型SOMに基づいた大規模月地震波形の可視化システム. 第6回データ工学と情報マネジメントに関するフォーラム, E1-4 (2014).
- [6] P. Cudre-Mauroux, H. Kimura, K. -T. Lim, J. Rogers, S. Madden, M. Stonebraker, S. B. Zdonik, and P. G. Brown: SS-DB: A Standard Science DBMS Benchmark. http://www-conf.slac.stanford.edu/xldb10/docs/ssdb_benchmark.pdf