



SETI@home 一分散処理による地球外生命体探索一

立川 崇之

早稲田大学大学院理工学研究科物理学及応用物理学専攻 〒169-8555 新宿区大久保 3-4-1
e-mail: tatekawa@gravity.phys.waseda.ac.jp

SETI@home 計画は観測された電波を世界中のパソコンで分散処理し、地球外生命体からの信号を捜し出そうという計画である。1999年4月の開始以来、主催者の予想を大きく上回る100万人以上の参加者を集めている。ここでは主にパソコンでどのような処理がなされるか、また計画進行中に発生した問題点や、本来の目的以外で注目されている事柄などを紹介する。

1. はじめに

先日話題になった映画「コンタクト」は主人公が地球外の生命体からのメッセージを電波で受け、それに答えて接触を試みるというストーリーであった。現実人類がこのように地球外の電波の交信を確認する事はないだろうか。

SETI プロジェクトは知的生命体からの電波を探索する計画である¹⁾。ただ電波といっても地球上での電波による交信を考えてみれば、様々な周波数による通信が考えられる。テレビ放送とは違い、交信のために発信された電波はいつどの方向からやってくるかわからない。そのために受信した膨大な電波のデータを解析する必要がある。

非常に高価なスーパーコンピューターを使う代わりに、世界中の沢山のパソコンを使って分担してデータ解析を行おうという計画がSETI@homeである^{2), 3)} (図1)。このような分散処理の計画は、爆発的に普及したインターネットがあつて初めて可能になるものである。

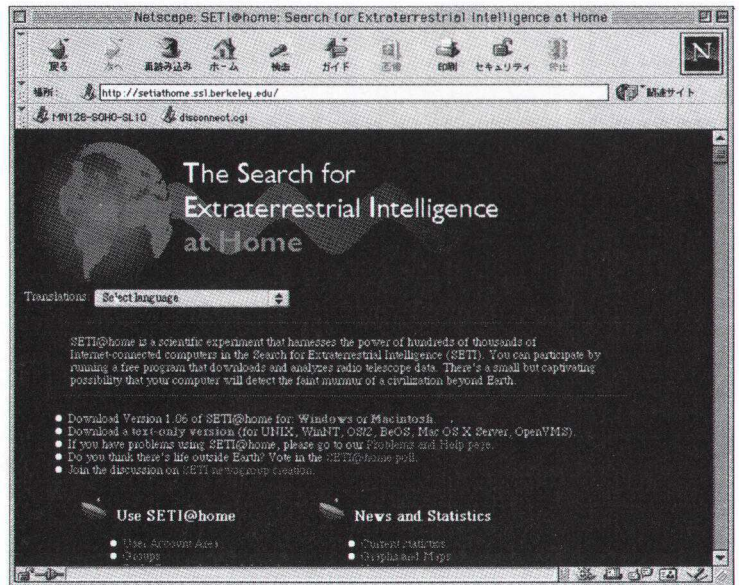


図1 SETI@homeのウェブページ。解析結果、ニュース、最新版のソフトの情報などが逐次更新されている。専門家向けの技術文書も置かれている。

世界中のパソコンを使った分担処理のプロジェクトは、SETI@homeが初めてではない。例えば暗号ソフトを開発しているアメリカのRSA社は、暗号を武器とみなして海外での長い暗号鍵を禁止した合衆国の方針に反論するため、暗号解読コンテストを行っている。合衆国の方針に則った暗号がいかに安易で破られやすいかを示

すために実施しており、その方法とは可能性のある全ての鍵を試して暗号が解読できるかどうかをチェックするというものである。可能性のある鍵が膨大な数にのぼっても世界中のコンピューターを駆使すれば破られるまでにさほどの時間はかからず、暗号の種類によっては24時間以内に破られた例もある⁴⁾。最近では円周率を求める計算など、様々な巨大分散計算プロジェクトが立ち上がっている⁵⁾。

SETI@homeはこのコンテストとは違い、答えつまり地球外生命体からの信号が解析によって見つかるという保証はないし、賞金がかかっているわけでもない。しかしプロジェクト開始以後、参加者は増え続けている。最近、主催者が参加理由を尋ねたアンケート結果が公表され、生命体を探すことに興味を持つ事や有名になる事以外にも様々な理由が挙げられている。例えば「新しい科学実験に参加するため」という純粋な興味からのものや、「我が宇宙で孤独ではないことを証明するため」、「ETからの信号は文明はその技術的な幼年期を生き延びることができるということの証明だ」という意味深なもの、「マイクロソフト製品をいくらかマシに使うため」「新しいオーバークロックさせたCPUを燃やすため」というコンピューターファンの側のもの、「ETの敵を見つけて人類を統一させるため」というSF小説の様なものもある。でもごく自然に「私にできる事だから」「なんでしないんだい？」と冷静に言ってしまう人もいる。

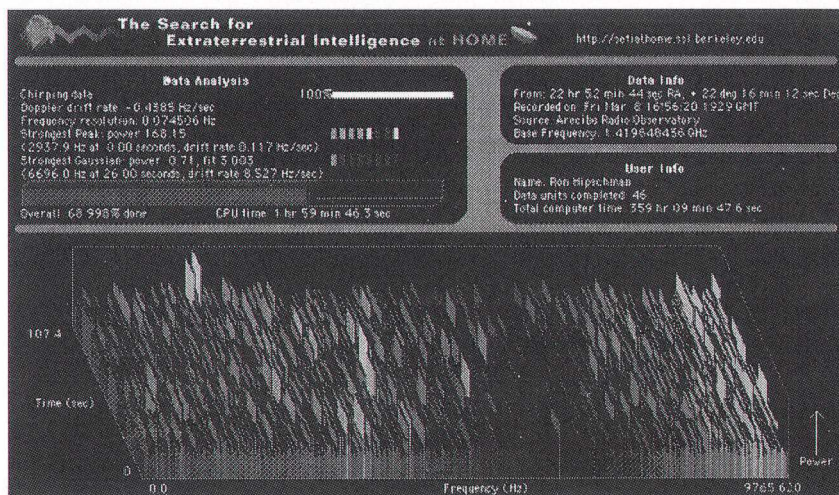


図2 クライアントソフトの解析中の画面。解析中のデータの電波強度のグラフ、データに関する情報、参加者の解析の統計などが表示される。

2. SETI@home のソフトはどのように動作するか

計画の一番の問題は、膨大なデータをどのように取り扱うかである。幸いにもこのデータはバラバラに並列に処理が行なえるように、細かい小片に分解できるようである。SETI計画では、プエルトリコのアレシボ天文台において2年間で全天を3回観測するそうだが、得られるデータは1日で35ギガバイト（1ギガバイト＝10の9乗バイト、英数字一字が1バイト、漢字一字が2バイト）にも及ぶ。このデータを磁気テープに記録しパークレーに郵送して、0.25メガバイト（1メガバイト＝10の6乗バイト）ずつのワークユニットと呼ばれる小片に分解する。分解の方法として1420MHzを中心とした2.5MHzの周波数帯の観測データを10KHzずつに分解し、107秒分ずつに区切る。ワークユニットには付加情報を加えるため、一つがおよそ340キロバイト（1キロバイト＝10の3乗バイト）になる。約90キロバイトの付加情報にはデータを観測した時刻、方向や他のワークユニットと共通する電波信号のデータが含まれ



る他、後述するソフトの誤動作、悪意ある参加者による報告データの改変を見抜くような仕掛けがなされている。

参加者は予めクライアントソフトを入手しておく。このソフトはスクリーンセーバーになっており、参加者がパソコンを日常業務に使用していない時に起動し、再び操作を開始したら停止するようになっている(図2)。しかし前述の参加理由の欄を見ると、SETI@homeのためにパソコンを24時間稼働させているという人もいたようである。ソフトはデータを入手する際に参加者にインターネットへの接続の許可を求めようになっており、許可すればバークレーのサーバーからデータの受信を行い、処理が終われば解析結果を送信する。データの受信は家庭からモデムを使用して行っても、せいぜい数分で済むような大きさになっている。

解析において問題となるのは、地球人起源の電波源である。このような人工的な電波源は強度を長時間持続する。一方、天体が電波望遠鏡で捕らえられる時間はわずか12秒で、電波強度はガウシアンの変化する。ソフトでは強度の強いガウシアンの変化を探し出すと試みる。またこのソフトでは生命体が効率的な送信をしていると仮定し、いくつかの狭いバンド幅で解析を行う。さらに、地球と他の惑星系との相対運動によるドップラーシフトも調査する。

12秒間でガウシアンの変化する電波源を見つけたら、ソフトはバークレーのSETI@homeチームへ結果を送信する。ここでパルサーや活動銀河核などのような既知の電波源のデータベースと比較し、地球外生命体からの信号である可能性のないものを排除する。排除されるデータは検出されたものの実に99.9999%になるという。もし排除されなければ、同じ対象物を別の観測と照らし合わせる。さらなる観測でも信号が検出され、しかも排除されるべきデータでなければ、SETI@homeチームは他のグループに調査を依頼

する。これは新彗星発見の際の動きに似ている。他のグループによる調査によって、コンピュータやソフトのバグなどを除去できる。非常に離れた場所での干渉測定により、対象物の距離もわかるようになる。これで確認されたら、スクリーンセーバーを動かして信号を検出した人はSETI@homeチームとともに、共同発見者の一人として名が記され、榮譽を受けるというわけである。

3. SETI@home 計画進行上の問題点

全世界のパソコンを利用して受信データを解析するプロジェクトが開始したが、道のりは平坦ではなかった。計画者達はSETI@homeの参加者は五万人程度であると想定していた。しかし1999年9月現在、参加者は百万人を越えている。このため、データの送受信を行い集計をとるためのサーバーが過負荷になってしまい、クライアントソフトの要求に応えきれなくなってきた。そのため、処理済みのデータの返送や新たなワークユニットの取得に支障が起き、しばしば失敗する事がある。予想外の参加者の多さに、主催者側ではスポンサーを頼ってサーバー設備をかなり増強したようである。だがそれ以上の参加者の増加、データ処理量および速度の上昇についていけなくなる事もしばしばある。

データはプエルトリコから磁気テープに記録され、バークレーに運ばれる。このデータを分割して配付するのだが、参加者の需要に追い付かなくなった際にはとんでもない事態に陥る。参加者の要求は常に発生するが、配付するべき未解析のデータが存在しない。この時に解析済のデータを再び配付してしまうのである。日本のメーリングリストのやり取りでは、1999年1月上旬のデータばかり解析していて、それ以降のデータを解析している人がいない時期があった。しばらくして主催者からデータの配付に関するお詫びが出ていた。参加者の意欲をなくさないように、このような事態は避けなければならないだろう。現在の観

測では一日に14万のワークユニットが生成される計算になる。一般的なパソコンで一つのワークユニットの解析に24時間程度かかるといわれているので、全員が一日の1/4だけ動かしたらデータの供給が追いつかなくなっていく計算になる。特に最近のパソコンでは処理が速い。トラブルでの処理の遅れは十分にカバーできるほどの参加者を集めているが、処理すべきデータがない場合は解析済のデータを再配布するのでは

なく、ソフト側で処理すべきデータがないというメッセージを表示するようになっていけば便利である。

この計画はテーマと全世界で行われることから注目を集めるようになった。このためにクラッカーのいたずらにも遭うようになっていた。SETI@homeのウェブページがたわいもないいたずら書きにすり替えられた事が一回あった。この際は主催者、参加者共に冷静に対処し、マスコミが大きく取り上げる事もなく、また主催者達も信頼を損なう事なく直ちに復旧された。また、クライアントソフトでは参加者の所属する国や地域をメニューから選ぶようになっており、しばしば不思議な国における解析の報告がなされているが、ここを改造して「火星」や「ケンタウルス座α星」などと報告する人も出た。もちろん嘘の報告である。

今までになされた中でもっとも悪質と思われるのが、解析が終わっていないデータを「解析済」として返送する事である。後述するが、解析数と解析速度はウェブページで公開される。ここでは様々な分類で解析数のランキングを行っており、上位に入ろうという目的でデータを改ざんする。

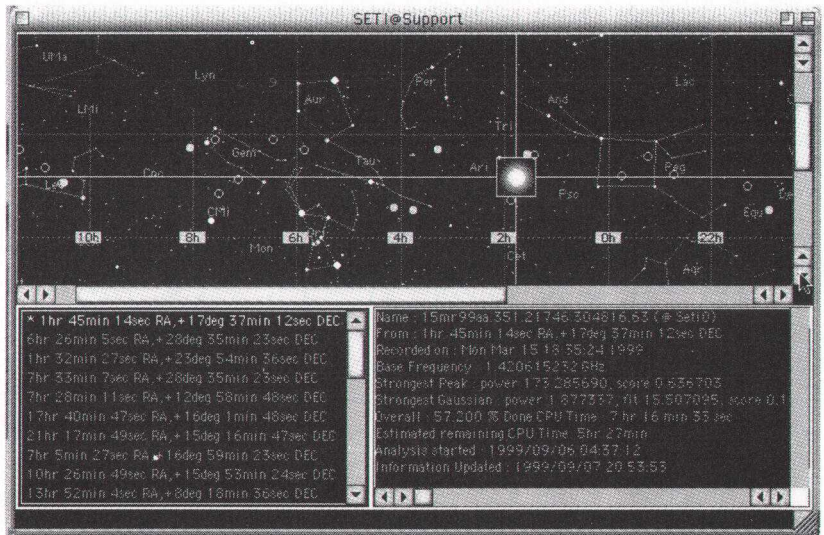


図3 解析状況表示プログラム SETI@Support の表示画面。解析中のデータの天球のどの位置に相当するかが表示される。

一時期とんでもない量の解析を行ったという報告があったが、前述した付加情報のおかげでいたずらを見抜く事が出来たようである。

悪質な例ではなく、一部のバージョンのソフトにおいてはデータのごく一部を解析しただけで、「解析済」という結果を返してしまうバグがあった。これも付加情報のおかげで見抜く事が出来たようで、ソフトのバグ修正とともにデータベースも修正された。もちろん、付加情報の中身や解析済でないデータをチェックするアルゴリズムは公表されていない。公表すれば、その裏を突かれる恐れがあるからである。

4. 予想外の方向へ進展した SETI@home 計画

もちろん、SETI@homeに関わるコンピューターが得意な人が悪い人ばかりというわけではない。まだ対応したソフトが公開されていないプラットフォームへのソフトの移植を行った人たちがいる。様々なUNIXへの移植がなされたが、その中でも極端な例を挙げると、SETI@homeを通じて

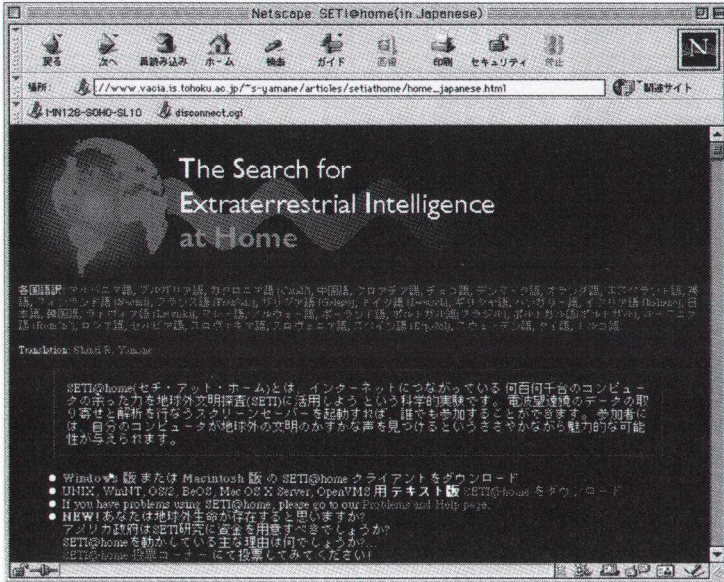


図4 SETI@homeの日本語ウェブページ。図1で示した英語版のものと同じ構成になっている。一部英語のままのものがあるが、ボランティアによって日本語の訳がなされている。

自分達の開発した無料OSのSCOを一緒に宣伝する場合⁶⁾や、社員によって次々と機能が付け加えられているCompaqのOSの例⁷⁾がある。さらに現在でも、いくつかのOS対応のものが開発中となっている。

ソフトの移植どころか、オリジナルのクライアントソフトに様々な意味を持たせ、参加者達の意欲を増すようなソフトウェアやツールを作っている人たちがいる。たとえばオリジナルのソフトでは解析しているデータの情報、すなわち対象としているデータの観測時刻、方向は座標の数値でしか表示されない。これではどの星座のあたりを解析しているのかがすぐには分からない。そこで、自分が解析しているデータが天球上のどのあたりのものなのかを示してくれるツールを作った人がいる⁸⁾(図3)。他にも様々なツールを作成している人が世界中にいるようで、主催者以外にも計画を盛り上げてくれるボランティアが数多く参加している事が分かる⁹⁾。

コンピューター以外の面でもボランティアは活

躍している。例えばSETI@homeのウェブページは英語版しかなかった。これを各国語に訳しているのはボランティア達である。ボランティアが計画やソフトの動作に関する概要を訳しているため、英語の知識がそれほどなくても簡単にソフトを入手して参加できるようになっている¹⁰⁾(図4)。主催者が発表したニュースについても、ただちにメーリングリストやウェブページで訳されたものを見る事ができる。サーバーなどのシステムにはスポンサーがついているが、参加者側は数多くのボランティアが活動している。日本だけでも数万人が参加しているのは、このボランティアの活動が影響しているかも知

れない。ウェブページには電波天文学に関する専門的な文書も含まれているので、電波天文を専門としている方に翻訳ボランティアとして参加していただけたらとてもありがたい。

ボランティアでなくても、メーリングリストで積極的に発言することによってお互いに得るのは大きい。日本語でやり取りされている参加者のメーリングリスト¹¹⁾の内容を見ると、実に様々な人々がそれぞれの立場から発言しており興味深い。天文ファン、SFファン、コンピューターやネットワークに詳しい技術者などがいろいろな情報を提供しており、ソフトを使用している際のトラブル処理だけではなく、関連事項についてもやりとりの間にいろいろと知る事ができる。

SETI@homeは知的生命体探索のためにソフトを実行しているが、この処理速度の比較から逆にCPUの比較をすることもできる。RSA社の暗号解読でも同様の比較がなされているが、暗号解読は素因数分解が主なアルゴリズムであるのに対し、SETI@homeは高速フーリエ変換が主なアル

ゴリズムである。よって浮動小数点計算を多用するため、理論天文学における数値計算のためのCPU 選択に関して参考になるかもしれない。主催者のサイトでは大まかな分類でCPU のタイプと所要時間が公開されている。Intel 系と PowerPC では同レベルのように見える。詳細な比較は日本のユーザーによってなされている。浮動小数点計算に強い Alpha ならばかなりの性能を発揮できると考えられ、結果はやはりそのようになっている。最新機種に純正の UNIX を搭載すれば、わずか一時間でワークユニットの解析が終わってしまうようである。1999 年 9 月に発表されたアップル社の新しいパソコンの発表会では、性能を示すためのデモに SETI@home のクライアントが使用された。ここまでくると、このプログラムはパソコンの性能を比較するためのベンチマークテストとして標準的なものになったのかと考えられるようになる。

5. おわりに

計画終了まではまだ一年以上の期間がある。今から参加しても十分に間に合う。知的生命体の存在確率を示すドレーク方程式¹²⁾において、パラメーターの選択によっては二年間の計画中に信号を捕らえる可能性が十分高くなる。期間中に「もしや?」と思われる信号が検出できるかも知れない。その際には主催者の提案のように冷静に対処しなければならない。もし信号が検出されなくても、100 万人以上が地球規模で参加したこの巨大プロジェクトは、今後の天文学の発展に大きな意義を持たせる事になると思われる。

本原稿の執筆に際し、私を執筆者として推薦し下書きのチェックを行った東北大学の山根信二氏、執筆の提案をされた天文月報の内藤統也氏に感謝の意を表する。

参考文献

- 1) Cordes J. M., Lazio T. J. W. and Sagan C., 1997, ApJ 487, 782-808
- 2) <http://setiathome.ssl.berkeley.edu/>
- 3) <http://setiathome.ssl.berkeley.edu/sciencepaper.html>
- 4) <http://www.rsasecurity.com/rsalabs/des3/index.html>
- 5) <http://www.cecm.sfu.ca/projects/pihex/pihex.html>
- 6) <http://www.sco.com/seti/>
- 7) <http://setiathome.ssl.berkeley.edu/README.Open-VMS.txt>
- 8) <http://member.nifty.ne.jp/GOTsubo/contents/SETISupport/download.html>
- 9) <http://setiathome.ssl.berkeley.edu/links.html>
- 10) <http://www.vacia.is.tohoku.ac.jp/~s-yamane/articles/setiathome/>
- 11) <http://www.vacia.is.tohoku.ac.jp/~s-yamane/articles/setiathome/maillinglist.html>
- 12) セーガン C. (木村繁 訳), 1980, COSMOS (朝日新聞社) 下, 254

SETI@home -Search for Extraterrestrial Intelligence With Computers Scattered Around The World-

Takayuki TATEKAWA

Department of Physics, Waseda University, 3-4-1 Ohkubo, Shinjuku, Tokyo 169-8555

Abstract: SETI@home project was planned that data analysis was accomplished using distributed computing with the help of hundreds of thousands of participants and their internet connected computers. Since the project was started on April 1999, more than one million people had become participants. In this article, "how SETI@home works" and interested topics are mainly shown.