

# データを隠しながら共用する、ベイズ最適化による材料探索手法

## [1] 要旨

様々な物性値や材料特性の最適化にデータ科学的手法が広く使われるようになってきている。そのために、データを持ち寄って探索をより速めたいが、それぞれはデータを開示することには消極的である。別の選択肢としてデータを隠したまま共用を実現する秘匿計算技術がある。今回、この技術を用いたベイズ最適化による材料探索手法が報告された。この手法により、性能の良くないデータを共用する場合でも、少ない探索回数で最適化が行われることが統計的な性能評価によって示された。

## [2] 本文

近年データ科学的手法が物理学の分野でもさかんに使われるようになってきている。物性値や材料特性の最適化への適用はそのような例であり、これを発展させる形で材料の機能要求に応じた新材料を探索するデータ駆動型材料探索技術の発展が注目されている。ここで重要となるのは材料特性のデータである。これまでに蓄積されたデータであるが、データ科学としてすぐに利用できるように開示されたものとなると非常に限られている。秘匿計算技術はデータの開示はしないが共用は認めるという別の選択肢を提供する。

最近、産業技術総合研究所機能材料コンピュータショナルデザイン研究センターの研究グループが秘密分散に基づく秘匿計算技術を採用し、真の物性値を参照することなくベイズ最適化と呼ばれる方法を実行する材料探索手法の開発に成功した。さらに磁石材料データを用いた統計的な性能評価によってこの手法が実用的にも有力であることを報告した。この成果は JPSJ の 2025 年 1 月号に掲載された。

ここでは秘密分散に基づく秘匿計算技術が採用された。秘密分散とは隠しておきたいデータを無意味化して分散させることで秘密を守る方法である。例えば、ある物性データの値「88184」があるとしよう。これを2つの数の和に分解する方法は無数にあるから、 $88184=13871+74313$  と分解した時、それぞれの数にもはや意味は無い。これらの無意味化したデータをシェアと呼ぶ。このシェアを2つの信頼できるサーバーに分散させて保管すれば、この二者から復元する権限が与えられない限りは元に戻せない。このような秘密分散に基づく秘匿計算では、シェアのまま復元せずにデータの和や積を計算できる方法が知られている。これらを組み合わせて様々な種類の秘匿計算を遂行する事が原理的には可能だが、実装上は計算量や通信量などの問題を解決する必要がある。

本研究では、材料探索の手法として有力と考えられていたベイズ最適化の計算に秘匿計算技術を適用した材料探索手法を実装したことにある。秘匿計算におけるベイズ最適化のアルゴリズムについてはこれまでも議論されていたが、実システムの報告は先例が無かった。秘匿計算は従来より多くの計算資源と通信が必要となるため、実用的な材料探索の実装には実際のタスクの要所を押さえた近似手法の採用をすることによって高速化するなどの工夫が必要であった。

この研究では、132 個の磁石化合物の候補から、できるだけ少ない回数のデータ取得によって磁化が最大のものを探すという課題を設定し性能評価を行った。ここでは秘匿計算による新しい方法が、探索回数の点では通常のベイズ最適化と遜色のない効率を持つことが示された。通常の秘匿化なしの計算と比べると 1500 倍ほどの時間を必要とするが、1 回の候補選出は 5 分ほどで完了した。

また本研究はデータの秘匿共用シナリオによる探索の加速についても報告している。このシナリオでは、最高磁化を含まないグループから6点のデータ共用を認める。そのアイデアはベイズ最適化の初期段階でモデルがある程度妥当なものに達するまでに必要な探索コストを、データ共用によって回避することにある。ベンチマークにおいて、秘匿データの共用により初期データ取得なしに速やかにデータの探索が完了したことが報告されている（図1）。磁化の低いデータの共用であっても探索効率は保たれており、最適化において不利なデータは探索範囲を実質的に狭める効果があるためであると議論されている。

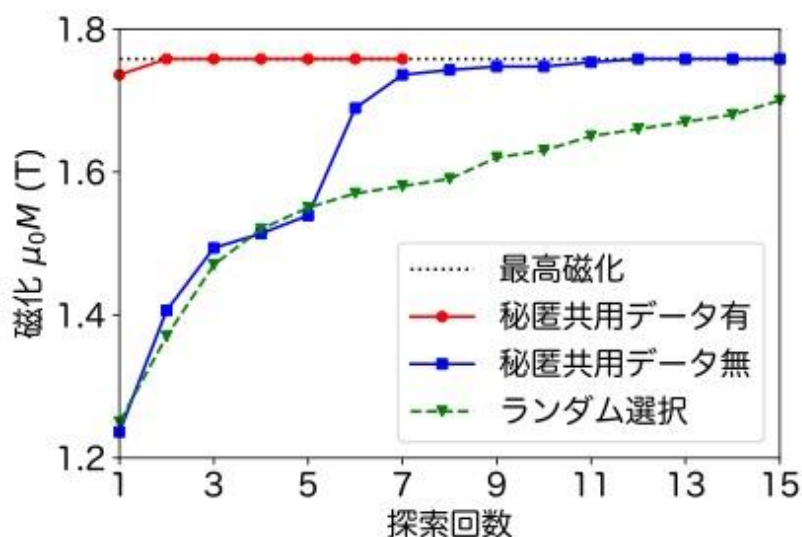


図1. 探索回数と磁化の関係。データ点は試行の90%がこれより高い磁化を持つ系を見つけたことを表す。

#### 原論文（2024年12月24日公開済）

Materials Secure Computation with Secret Sharing: A Bayesian Optimization Scheme and Its Performance

T. Fukazawa, T. Ikegami, M. Kawata, and T. Miyake, J. Phys. Soc. Jpn. **94**, 013801 (2025).

< 情報提供：深澤太郎（産業技術総合研究所 研究チーム付） >