

同時発表：
筑波研究学園都市記者会（資料配布）
文部科学記者会（資料配布）
科学記者会（資料配布）



Tokyo Tech

機械学習の「記憶」を活用し、高分子の熱伝導性の大幅な向上に成功

～少ないデータでも高精度な予測が可能に 高分子での材料インフォマティクス加速に期待～

配布日時：2019年6月26日14時

国立研究開発法人 物質・材料研究機構

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

国立大学法人 東京工業大学

概要

1. NIMS、統計数理研究所、東京工業大学の共同研究グループは、独自の機械学習の解析技術を用いて高熱伝導性高分子を設計・合成し、従来の高分子に比べて約80%の熱伝導率の向上に成功しました。同グループは、少数の物性データから予測モデルを導くために、転移学習^①と呼ばれる解析技術を駆使して問題解決を図りました。今回の成果は、新材料の発見のみならず、高分子インフォマティクスの最大障壁とされる「スモールデータ問題」の克服に向けた大きな一歩と位置付けられます。

2. 一般に高分子の熱伝導率は金属やセラミックスに比べて非常に低いことが知られています。一方で近年の高分子研究により、特異的に高い熱伝導率を持つ高分子が存在することが明らかになってきました。このような背景から、自動運転システムや次世代無線通信規格5G等、放熱性の向上が求められるエレクトロニクスデバイスの開発において、成形性に優れた高分子材料の高熱伝導化の研究に注目が集まっています。

3. 同グループは、世界最大の高分子データベース PoLyInfo^②と独自の機械学習アルゴリズムを組み合わせ、高熱伝導性を持つ新規高分子の設計に取り組みました。PoLyInfoには、ホモポリマーに限定した場合に、室温付近の熱伝導率のデータが28件（種類）しか登録されていません。そこで、ビッグデータの入手が可能な他の物性データ（ガラス転移温度等）で機械学習のモデルを訓練し、モデルが獲得した「記憶」と少数の熱伝導率のデータを組み合わせることで、熱伝導率を高精度に予測できるモデルを導くことに成功しました。これは一般に転移学習と呼ばれる解析技術です。同グループは、このモデルを用いて高い熱伝導率をターゲットに1,000種類の仮想ライブラリ^③を作製しました。その中から三種類の芳香族ポリアミド^④を合成し、熱伝導率0.41 W/mKに達する高分子を見つけ出しました。これは、典型的なポリアミド系高分子（無配向）と比較して最大80%の性能向上に相当します。さらに、同グループが開発した材料は、高耐熱性や有機溶媒への溶解性、フィルム加工の容易性等、実用化のステージで求められる複数の要求特性を併せ持つことも実験的に確認されました。

4. 一般に、材料データは取得コストが高く、情報漏洩の観点から研究者にはデータの公開に対するインセンティブが働かないため、材料インフォマティクスのデータ量は、少なくとも短中期的には、大学のラボや企業で生産可能な水準に留まることが予想されます。同グループが開発した転移学習の解析技術は、材料インフォマティクスのスモールデータ問題の克服に大きく寄与することが期待されます。

5. 本研究は、情報・システム研究機構 統計数理研究所 ものづくりデータ科学研究センター 吉田亮教授（同センター・センター長）と Wu Stephen 助教、東京工業大学 物質理工学院 森川淳子教授、物質・材料研究機構 統合型材料開発・情報基盤部門 情報統合型物質・材料研究拠点 伝熱制御・熱電材料グループ 徐一斌グループリーダーらによって行われました。また本研究は、科学技術振興機構（JST）のイノベーションハブ構築支援事業「情報統合型物質・材料開発イニシアティブ（MIPI：“Materials research by Information Integration” Initiative）」（法人名：物質・材料研究機構、プロジェクト実施期間：2015-2019年度）の支援の下で推進されました。

6. 本研究成果は、英国時間2019年6月21日午前10時（日本時間21日午後6時）にnpj Computational Materials誌にて発表されました。

研究の背景

一般に材料設計のパラメータ空間は極めて広大です。例えば、有機化合物のケミカルスペースには、 10^{60} を超える候補物質が存在すると言われています。さらに、実材料の開発では、プロセス、添加剤・溶媒選択、膜材料の層構成等の制御因子が加わり、パラメータ空間の次元は爆発的に増大します。材料インフォマティクスの多くの問題は、このような広大な探索空間から所望の特性を有する埋蔵物質を発掘することに帰着します。

同グループが開発した iQSPR[®]は、所望の特性を持つ化学構造を設計する機械学習アルゴリズムです。高分子データベース PoLyInfo の実験データを iQSPR に入力し、高熱伝導率をターゲットに候補物質の仮想ライブラリを構築しました。さらに、三種類の芳香族ポリアミドを選定・合成し、熱伝導率 0.41 W/mK を達成する新しい高分子を発見しました (図 1 参照)。

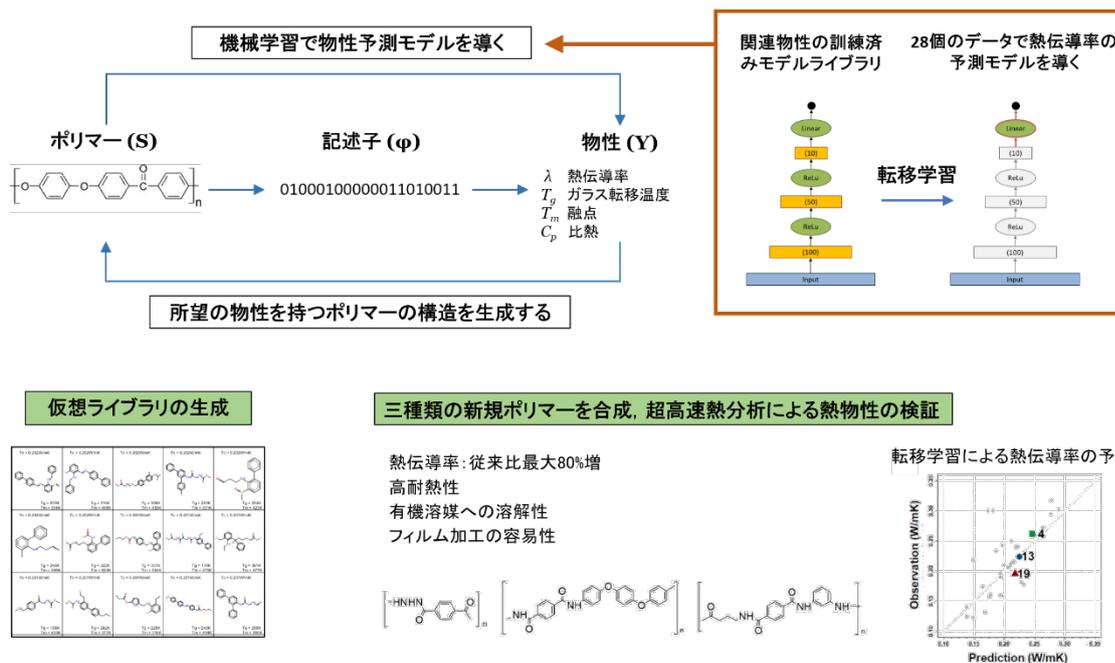


図 1: 三種類の高熱伝導性高分子の発見に至るワークフロー。転移学習を活用した熱伝導率の予測と分子設計の機械学習の技術が問題解決の突破口を切り拓いた。

研究内容と成果

iQSPR のワークフローは、順方向と逆方向の計算から構成されます。機械学習でポリマーの構造から特性の順方向の予測モデルを構築し、その逆写像を求めることで、特性から構造の逆方向の予測モデルを導きます。このモデルを用いて仮想ライブラリを作成し、所望の特性を有する埋蔵物質を発掘します。しかしながら、PoLyInfo に登録されている熱伝導率のデータはたったの 28 件しかなかったため、従来の機械学習では物性予測のモデルを作成することができませんでした。

そこで同グループは、転移学習という解析技術を導入して問題解決を図りました。まずは、ビッグデータが入手できる他の物性に関するデータ (高分子のガラス転移温度、低分子化合物の比熱容量等) を収集し、機械学習のモデルライブラリを構築しました。データに基づく構造・物性の学習を経ることで、これらのモデルは高分子の構造に関する「汎用的な内部表現」を獲得しました。このように「経験」から獲得した「機械の記憶」を適切に活用することで、たった 28 件の熱伝導率のデータでも十分な精度を達成する予測モデルを得ることができました。優れた研究者は、過去の経験から大量かつ多様な知識の体系を構築し、データがほとんど存在しないような新しいタスクに対しても合理的に予測や意思決定を行うことができます。同グループが開発した転移学習のアルゴリズムは、まるで熟練の材料研究者の認識・判断の過程を模倣したかのようなパフォーマンスを発揮しました。

同グループは、このような解析技術を用いて、高熱伝導率をターゲットに 1,000 種類の高分子の仮想ライブラリを設計しました。その中から三種類の芳香族ポリアミドを合成し、最大で熱伝導率 0.41 W/mK に

達する高分子を発見しました。また、実験結果は機械学習の予測とほぼ一致しました。同グループが達成した熱伝導率は、典型的なポリアミド系高分子（無配向）と比較して約 80%の性能向上に相当します。さらに、高耐熱性や有機溶媒への溶解性、フィルム加工の容易性等、今度の実用化フェーズで重要になる諸特性を併せ持つことが実験的に確認されました。また、従来の熱分析技術では高耐熱性高分子のガラス転移温度を測定できなかったため、最新の超高速熱分析技術を新たに開発し、高温域の転移温度の測定に成功しました。

今後の展開

本研究は、機械学習が自律的に設計した高分子が実際に合成・検証された初の事例となります。近年、材料研究とデータ科学の融合が急速に進行し、その有効性や可能性について、実証的見地から様々な検討が行われています。しかしながら、他の領域に比べると、高分子研究のデータ科学との学融合は大幅に遅延しています。その背景には、多くの高分子物性はデータ科学の解析手法を適用できるほどのデータ量に達していないという自明な理由が存在します。今後、高分子インフォマティクスでは、スモールデータの限界をいかに突破するかが勝利の鍵を握ります。同グループの成果は、当該分野が抱える本質的な問題の克服に一石を投じるものです。

また、今回は合成の容易性という観点から三種類の高分子を選定・合成しましたが、仮想ライブラリには他にも有望な候補物質が数多く残されている可能性があります。また、同グループが開発した機械学習の技術は汎用的なものであり、任意の特性をターゲットに同様の解析を行うことができます。これから数年以内に、同じようなアプローチで多くの埋蔵物質が発掘され、その中から、従来の常識を覆すような新しい高分子材料が発掘されることが期待されます。

掲載論文

題目：Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm

著者：Stephen Wu, Yukiko Kondo, Masaaki Kakimoto, Bin Yang, Hironao Yamada, Isao Kuwajima, Guillaume Lambert, Kenta Hongo, Yibin Xu, Junichiro Shiomi, Christoph Schick, Junko Morikawa, Ryo Yoshida

雑誌：npj Computational Materials

掲載日時：2019年6月21日午後6時（英国夏時間21日午前10時）

用語解説

- (1) 転移学習：あるタスクの学習モデルを別のタスクに流用することを目的とする方法論の総称。例えば、膨大なデータから訓練された動物の種類を判定する画像認識の多層ニューラルネットワークを改変し、少数の花の画像データを用いて分類器を構築したいと考えます。動物の分類器は、学習過程で画像認識に必要な基本的な特徴量を抽出していることが期待され、その中の一部は花の分類にも流用可能であると考えられます。その場合、花の分類器を一から学習するのではなく、少数のデータを使って動物の分類器を微修正すれば十分かもしれません。このような推論アルゴリズムの総称が転移学習です。転移学習という用語はさらに広い概念を含みますが、とりわけスモールデータ問題に対する有効なアプローチであることが知られています。
- (2) PoLyInfo：国立研究開発法人 物質・材料研究機構が保有する高分子物性の世界最大級のデータベース (<https://polymer.nims.go.jp/>)。学術文献から収集した約 100 種類の物性（熱物性、電気的特性、力学的特性等）、化学構造、測定条件、重合方法等を収録しています。
- (3) 仮想ライブラリ：特定の用途をターゲットに計算機で作製した仮想物質のプール。機械学習の物性予測モデルと組み合わせ、所望の特性を持つ新規物質の候補を絞り込む際に使用されます（一般に仮想スクリーニングと呼ばれる）。同グループは、機械学習で熱伝導率や耐熱性をターゲットに 1,000 個の仮想高分子を作製しました。
- (4) 芳香族ポリアミド：ポリアミドは、主鎖に酸アミド結合（-CO-NH-）を持つ高分子の総称です。主

鎖にベンゼン核を有するポリアミドを芳香族ポリアミドといい、中でも、全芳香族ポリアミド(アラミド)はエンジニアリング・プラスチックとして、優れた耐熱性と強度を持つことが知られています。

- (5) iQSPR : 同グループ吉田らが開発した分子設計の機械学習アルゴリズム (Ikebata, H., Hongo, K., Isomura, T., Maezono, R. and Yoshida, R. (2017). Bayesian molecular design with a chemical language model, *Journal of Computer-Aided Molecular Design*, 31(4), 379-391)。実験やシミュレーションから得られるデータを用いて、物質の構造から物性の順方向の予測モデルを構築し、物性から構造の逆写像を求めて仮説物質を発生させ、所望の物性を有する埋蔵物質を炙り出すものです。確率的言語モデルに基づく構造生成器や機械学習の様々な解析技術を駆使して開発した確率推論のアルゴリズムです。

本件に関するお問い合わせ先

(研究内容に関すること)

国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門 情報統合型物質・材料研究拠点 物質・材料記述基盤グループ グループリーダー (大学共同利用機関法人 情報・システム研究機構 統計数理研究所 ものづくりデータ科学研究センター 教授・センター長)

吉田 亮 (よしだ りょう)

TEL: 050-5533-8534

E-mail: yoshidar@ism.ac.jp

URL: <http://daweb.ism.ac.jp/yoshidalab/>

国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門 情報統合型物質・材料研究拠点 伝熱制御・熱電材料グループ 特別研究員 (東京工業大学 物質理工学院 教授)

森川 淳子 (もりかわ じゅんこ)

E-mail: morikawa.j.aa@m.titech.ac.jp

(報道・広報に関すること)

国立研究開発法人 物質・材料研究機構 経営企画部門 広報室

〒305-0047 茨城県つくば市千現 1-2-1

TEL: 029-859-2026, FAX: 029-859-2017

E-mail: pressrelease@ml.nims.go.jp

大学共同利用機関法人 情報・システム研究機構 統計数理研究所 運営企画本部企画室 URA ステーション
〒190-8562 東京都立川市緑町 10-3

TEL : 050-5533-8580, FAX : 041-526-4348

Email: ask-ura@ism.ac.jp

国立大学法人 東京工業大学 広報・社会連携本部 広報・地域連携部門

〒152-8550 東京都目黒区大岡山 2-12-1

TEL: 03-5734-2975, FAX: 03-5734-3661

Email: media@jim.titech.ac.jp