

解禁時間 (テレビ、ラジオ、インターネット) : 令和元年6月6日 (木) 午前0時  
(新聞) : 令和元年6月6日 (木) 付朝刊

令和元年 6 月 3 日

報道関係者各位

国立大学法人 奈良先端科学技術大学院大学

## 生物の放出する物質と生物同士の複雑な関係を読み解くデータサイエンスの方法を開発 ～食材の栄養成分と味、香りなどの関係解明にも期待～

### 【概要】

奈良先端科学技術大学院大学 (学長：横矢直和) 先端科学技術研究科 情報科学領域 計算システムズ生物学研究室の Md. Altaf-Ul-Amin (モハマド=アルタフ=ウル=アミン) 准教授のグループは、生物と物質など 2 種の異なった対象物のデータが集積した二部構造を持つ大規模なネットワークについて、例えば、生態系の中の多数の生物とその生物が放出する多様な揮発性生化学物質の関係を視覚化し、分析するための方法を考案しました。

この研究は「二部グラフ・クラスタリング法」というデータサイエンスの手法を使ったもので、世界的に高く評価され、情報科学分野の最高峰の国際論文誌 IEEE TRANSACTIONS に採録されました。昨今、データサイエンスが注目される中、さまざまな分野への応用が期待されています。例えば、生態系において様々な生物がお互いの誘引物質や忌避物質を放出することによって化学的に会話しているという関係性をネットワークとして理解したり、あるいは、世界中のさまざまな料理に含まれる食材やそこに含まれる栄養成分の配合と味、香りなどの関係性を体系的に理解したり、といった多くの要素が複雑に関わる現象を読み解くための分析に応用することが可能です。

このようなネットワーク解析をもとにしたアプローチは、いわゆるビッグデータ医療を始め、ビッグデータサイエンス、データ駆動型サイエンスと呼ばれる、社会の様々な領域のデータ同士の関係性から有用な情報をとりだすマイニング (データの採掘) のための手法として重要であり、今後さまざまな発見を導く可能性が開けました。

(発表者) Mohammad Bozrul Karim, Ming Huang, Naoaki Ono, Shigehiko KANAYA, Md. Altaf-Ul-Amin

(論文タイトル) BiClusO: A novel biclustering approach and its application to species-VOC relational data

(掲載誌) IEEE TRANSACTIONS on COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, Jun. 1st (Print ISSN: 1545-5963), the United States

つきましては、関係資料を配付するとともに、下記のとおり記者発表を行いますので、是非ともご出席くださいますよう、お願い申し上げます。

### 記

<日 時> 令和元年 6 月 5 日 (水) 午後 3 時～ (1 時間程度)

<場 所> 奈良先端科学技術大学院大学 附属図書館 マルチメディアホール (3 階)  
奈良県生駒市高山町 8916-5 (けいはんな学研都市)

※アクセスについては、<https://www.naist.jp/>をご覧ください。

<説明者> 奈良先端科学技術大学院大学 先端科学技術研究科

情報科学領域 計算システムズ生物学研究室

教授 金谷 重彦

准教授 Md. Altaf-Ul-Amin (モハマド=アルタフ=ウル=アミン)

博士後期課程3年 Mohammad Bozlul Karim (モハマド=バジルル=カリム)

【ご連絡事項】

(1)本件につきましては、奈良先端科学技術大学院大学から奈良県文化教育記者クラブをメインとし、学  
研都市記者クラブ、大阪科学・大学記者クラブへ同時にご連絡しております。

(2)取材希望がございましたら、恐れ入りますが下記までご連絡願います。

(3)記者発表に関する問い合わせ先

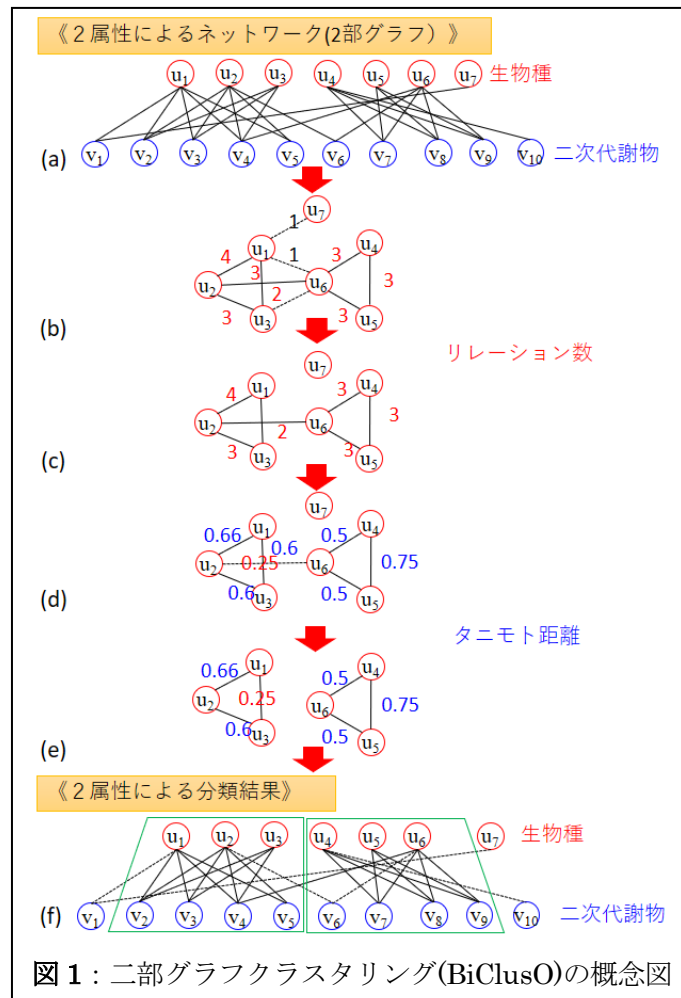
奈良先端科学技術大学院大学 企画総務課 広報渉外係 近藤 雄一郎 (こんどう ゆういちろう)

TEL: 0743-72-5026 FAX: 0743-72-5011 E-mail: s-kikaku@ad.naist.jp

【解説】

ビッグデータサイエンスの分野では、非常に多く要素の間の属性間の結びつきからなるネットワークを俯瞰的に把握し、その全体的な構造を理解することが大きな課題となっている。特に、様々な生物種 ( $u$ ) の中で生産される代謝物 ( $v$ ) の種類を結びつけたグラフ (図 1) のように、二つの属性同士の関係 (図 1.a) を考えた場合、その結びつきは全体として大きなネットワークとなる。しかし、そのネットワークの中で互いに結びつきが密な要素同士に着目すると、(図 1.f) のように小グループに分けられる。そのようなグループも効率よく見つけ出すアルゴリズム (計算手順) を、博士後期課程 3 年の学生、Mohammad Bozrul Karim (モハマド=バジルル=カリム) 君と、Md. Altaf-Ul-Amin (モハマド=アルタフ=ウル=アミン) 准教授が中心となって提案した。

アルゴリズムの概略を図 1 により説明する。それぞれの生物( $u$ )が代謝物( $v$ )を作るとき、その二つが結ばれているものとする (a)。まず生物種だけに着目し、共通する代謝物によって結ばれた生物種のネットワークをつくる (b)。始めにノイズを減らすため、共通する代謝物の数が少ないリンクを削除する (c)。得られたネットワークから、生成する物質の共通部分が多いもの同士が集まるように生物種同士の類似度を求め(d)、似た者同士のグループに分ける(e)。この手法は、代謝物のつながりから生物種をグループ分けすると同時に、共通に生産される代謝物をまとめる (f) ことができる効率的な探索方法と言える、また、生物と代謝産物の関係性の他にも、料理とその材料、ブログとそのトピックス、といったさまざまなジャンルのデータに応用することが可能である。



## 【今後の展開】

現在、ソフトウェアのプロトタイプ開発を進めており、活用したい研究者、データサイエンスに興味のある方々に無償にて提供したい。(図 2: BiClusO 解析アプリケーションのプロトタイプソフトウェア)

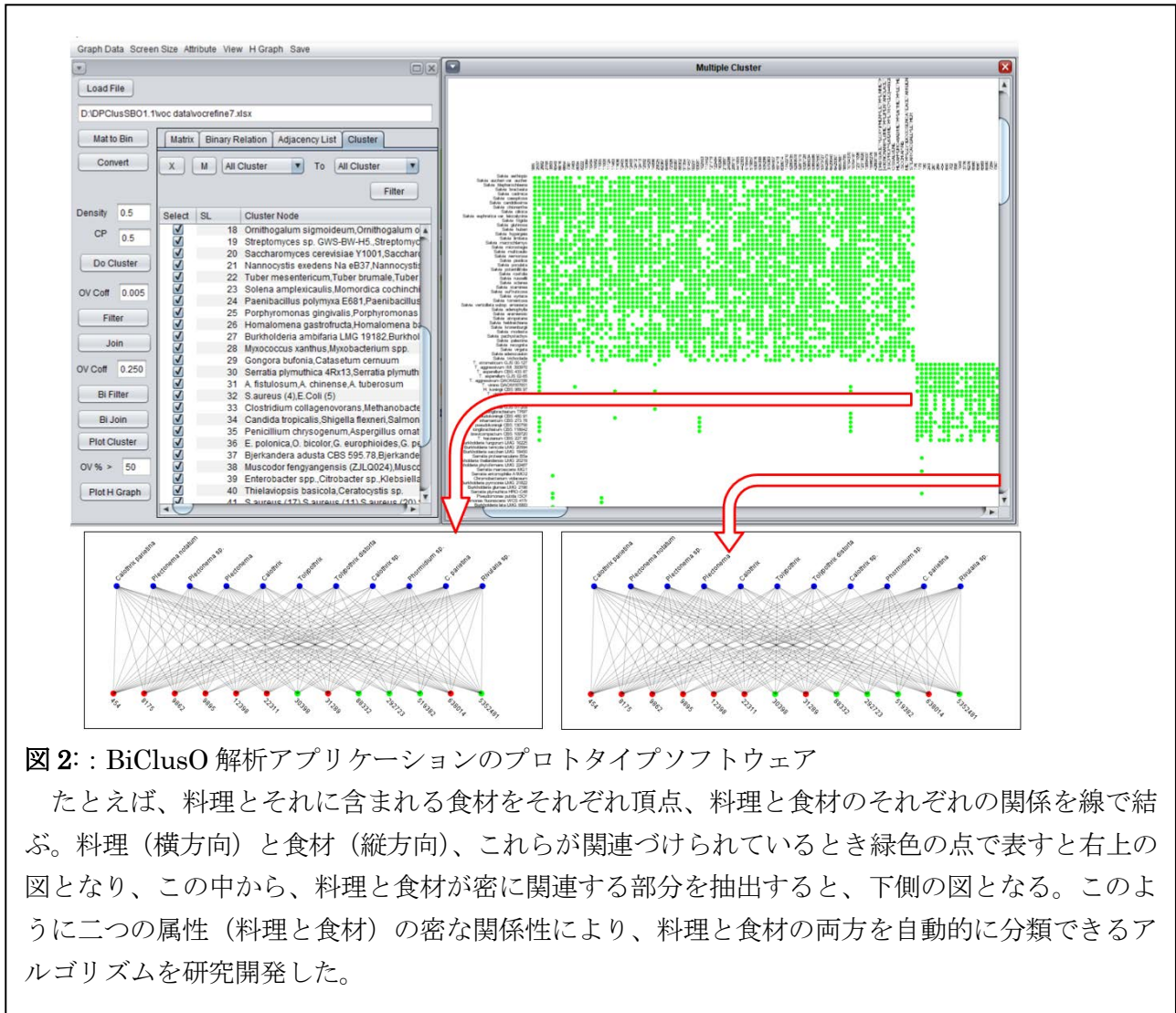


図 2: BiClusO 解析アプリケーションのプロトタイプソフトウェア

たとえば、料理とそれに含まれる食材をそれぞれ頂点、料理と食材のそれぞれの関係を線で結ぶ。料理（横方向）と食材（縦方向）、これらが関連づけられているとき緑色の点で表すと右上の図となり、この中から、料理と食材が密に関連する部分を抽出すると、下側の図となる。このように二つの属性（料理と食材）の密な関係性により、料理と食材の両方を自動的に分類できるアルゴリズムを研究開発した。

## 【用語解説】

### ○グラフ理論

グラフ理論とは、対象の集合を頂点集合、対象間の関係を辺集合からなるデータをもとに、対象間の関係性を理解するための数学・情報学

### ○二部グラフ・クラスタリング法

二つの異なる属性、例えば料理とそれに含まれる食材をそれぞれ頂点、料理と食材のそれぞれの関係を線で結んだグラフを二部グラフといい、その中で、お互いに密に関係性がある料理と食材を抽出する方法を二部グラフ・クラスタリングという。

【本研究内容についてコメント出来る方】

中村 哲 奈良先端科学技術大学院大学 データ駆動型サイエンス創造センター(DSC) センター長  
先端科学研究科 知能コミュニケーション研究室 教授

TEL : 0743-72-5952

E-mail : [s-nakamura@is.naist.jp](mailto:s-nakamura@is.naist.jp)

松本 健一 奈良先端科学技術大学院大学 先端科学技術研究科 ソフトウェア工学研究室 教授

TEL : 0743-72-5310

E-mail : [matumoto@is.naist.jp](mailto:matumoto@is.naist.jp)

【本プレスリリースに関するお問い合わせ先】

奈良先端科学技術大学院大学 先端科学研究科 計算システムズ生物学研究室  
教授 金谷 重彦

TEL : 0743-72-5952

E-mail : [skanaya@gtc.naist.jp](mailto:skanaya@gtc.naist.jp)

准教授 Md. Altaf-Ul-Amin (モハマド=アルタフ=ウル=アミン)

E-mail : [amin-m@is.naist.jp](mailto:amin-m@is.naist.jp)

,