

## Exact Clustering via Integer Programming and Maximum Satisfiability

宮内 敦史	菌部 知大	鮎川 矩義
Atsushi Miyauchi	Tomohiro Sonobe	Noriyoshi Sukegawa
理化学研究所	国立情報学研究所	東京理科大学

出典 : The 32nd AAI Conference on Artificial Intelligence (AAAI 2018), pp. 1387–1394

クラスタリングは、機械学習やデータマイニングで幅広く利用されているデータ解析手法である。クラスタリングをモデル化した最適化問題の一つとして、クリーク分割問題が知られている。この問題では、枝重み（正負あり）付き無向グラフが与えられ、頂点集合を任意の個数のクラスターに分けて、クラスター内の枝重みの総和を最大化することが要求される。クリーク分割問題に対しては、Grötschel & Wakabayashi (*Math. Program.* **45**, pp. 59–96, 1989) による整数線形計画問題としての定式化が知られており、整数計画ソルバーを用いて最適解を得ることができる。しかしながら、制約式の本数が非常に多く、実用上は数百頂点のグラフにしか適用することができない。本研究では、実インスタンスの性質に着目することで、制約式の本数を大幅に減らした定式化を提案する。提案定式化は、最大充足可能性問題としてのエンコーディングに書き直すことができ、MaxSAT ソルバーを用いて最適解を得ることもできる。計算機実験では、提案定式化と提案エンコーディングの性能を評価し、数千頂点のグラフに対しても適用可能であることを確認した。