

2024 年度

筑波大学情報学群情報科学類

卒業研究論文

題目

ソーシャルネットワークにおける多様な相互作用を  
捉えるエージェントベースモデル

主専攻 知能情報メディア主専攻

著者 岩橋七海

指導教員 岡瑞起

## 要 旨

この文書は筑波大学情報学群情報科学類の卒業研究論文のサンプルである。このファイルを書き換えて、このサンプルと同様の書式の論文を  $\text{\LaTeX}$  を使って作成できる。

このサンプルは、学生が論文を作成する手間を軽減するために提供している。このサンプルで示す書式はあくまで例であり、要項に準拠していれば、このファイルを使わずに自分で決めた書式を用いてもよい。

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 背景 . . . . .	1
<b>第2章 関連研究</b>	<b>3</b>
2.1 エージェントベースのポリアの壺モデル . . . . .	3
2.2 進化的アルゴリズム . . . . .	5
2.2.1 遺伝的アルゴリズム . . . . .	5
2.2.2 Quality Diversity アルゴリズム . . . . .	5
<b>第3章 手法</b>	<b>7</b>
3.1 戦略のベクトル表現 . . . . .	7
3.2 進化的アルゴリズムによる探索 . . . . .	7
3.3 評価指標 . . . . .	8
<b>第4章 実験</b>	<b>10</b>
4.1 データセット . . . . .	10
4.1.1 合成データ . . . . .	11
4.1.2 実データ . . . . .	11
4.2 パラメータ . . . . .	11
4.3 実験結果と考察 . . . . .	12
4.3.1 RQ1: 合成データによる精度比較 . . . . .	12
4.3.2 RQ2: 実データによる精度比較 . . . . .	15
4.3.3 RQ3: 実データにおける隣接可能空間の探索に用いられている戦略 . .	15
<b>第5章 結論</b>	<b>19</b>
謝辞	<b>20</b>

# 図目次

2.1	相互作用の3つの状態 . . . . .	4
3.1	戦略ベクトルを使った相互作用の例 . . . . .	8
4.1	パラメータ $(\rho, \nu, s) = (5, 15, SSW)$ の合成データと各手法の距離の箱ひげ図の比較 . . . . .	13
4.2	パラメータ $(\rho, \nu, s) = (5, 15, SSW)$ の合成データに対する GA と QD の探索の時間経過 . . . . .	14
4.3	パラメータ $(\rho, \nu, s) = (5, 15, SSW)$ の合成データに対する各測定指標の既存モデルと提案モデルの比較 . . . . .	15
4.4	実データと各手法の距離の箱ひげ図の比較 . . . . .	16
4.5	実データに対する GA と QD の探索の時間経過 . . . . .	16
4.6	実データに対する各測定指標の既存モデルと提案モデルの比較 . . . . .	17
4.7	ベクトル化した QD と GA アルゴリズムによる探索で見つけたネットワークの可視化 . . . . .	17

# 第1章 序論

## 1.1 背景

人間の社会は何気ない挨拶、取引や交渉、SNS でのコミュニケーションなどの様々な相互作用によって成り立っている。インターネットとモバイルデバイスの登場は、こうした多面的な相互作用に関するデータの記録と収集を容易にした。その中で、ソーシャルネットワークプラットフォームやオンラインソーシャルゲームを用いることで、人間の相互作用のモデリングや根拠にあるパターンや原理の発見が可能になる。これらのコミュニケーションを包括的に理解することは、新規ユーザーを獲得し、既存のユーザーを維持するための効果的なサービスの開発に貢献すると考えられる。

これまで人間の活動を再現し理解するために、数理モデルが用いられてきた。ネットワークは新たなノードやエッジの追加といった成長と、ネットワークの構造やエッジの重みといったダイナミクスの二つの側面から考えることができる。ネットワークの成長のモデルとしては、Barabasi-Albert(BA) モデルのような優先的選択性を捉え、スケールフリー性を表せるモデルがある<sup>1-3</sup>。BA モデルはネットワークの成長を表すことはできるが、ネットワークの完全なダイナミクスを表現することはできない。ネットワークのダイナミクスに焦点を当てたモデルは、現実世界のネットワーク構造をよりよく模倣するために、様々な手法を提案している。これらの方法には、1つのステップで複数のエッジを導入する方法<sup>4</sup>、友人が繋がりを作る可能性が高い triadic closure<sup>5</sup>、ユーザーの職業や信頼性に基づくフィットネスなどのノードの特性を考慮する方法<sup>6</sup>などがある。しかし、これらのモデルは、既存のノード間の関係を強化したり、新しい接続を形成したりするような活動をとらえることができず、複雑なネットワークの複雑さを完全に表現することはできない。

これらのモデルの原理を活用し、実際のデータを再現する精度を向上させるために、様々なモデルが提案されている。モデルの一つのカテゴリーは、新規性の出現は従来想定されていたようなランダムなものではなく、「隣接可能空間」から発生するという考え方を取り入れたものである<sup>7-9</sup>。隣接可能空間とは現在の観測範囲から1ステップ外の空間のことであり、何らかのきっかけが一つあれば観測できる要素によって張られている。Kaufman は当初、分子や生物の進化を解明するためにこの考えを提唱した<sup>10</sup>。同様の考え方はタンパク質空間説にも見られ、遺伝子の進化は既存の遺伝子の微小な変化の積み重ねによって起こるものとし、その変化は遺伝子が表現型を形成できる制約の下で起こらなければならないと主張した<sup>11</sup>。Kaufman はこの理論を遺伝子の進化だけでなく、人間関係の進化や技術革新の発展など様々な分野に拡張し、一般化した。

このような隣接可能空間の考え方は、Tria ら<sup>12</sup>によってモデル化された。このモデルはトリガー付き壺モデルと呼ばれ、環境に壺が一つ存在すると仮定し、壺からエージェントを表す玉を選択したり追加したりすることでシステムの動作を表す。Ubaldi らは、各エージェントにそれぞれの壺を割り当てることで、このモデルをエージェントベースのモデルに拡張した<sup>9</sup>。このモデルでは、隣接可能空間を通じて、エージェント同士のつながりを作ることで、社会的ネットワークが拡大すると仮定されている。エージェント個々の定義されたルールセットに従って、社会的相互作用のダイナミクスをカプセル化し、エージェントの行動を決定する。これらのルールは、ネットワークの成長とダイナミクスを捉え、既に相互作用したことがある人との相互作用、新しい人との相互作用、新しい人と関わるための隣接可能空間の探索を含む。比較的単純にもかかわらず、このモデルは Twitter 上で行われる返信や論文の共著、携帯電話の発信などの振る舞いを性格に表現できることが報告されている<sup>9</sup>。

しかし、隣接可能空間を探索するという考え方をモデルに組み込むことには課題もある。従来のモデルでは、人間がこの空間を探索する戦略をあらかじめ定義しておく必要がある。そのため、実際のデータに最も適合するパラメータの最適化探索は、これらの事前に定義した戦略に制約される<sup>8,9,13</sup>。この制限に対応するため、本研究では複数の戦略に対応し、これらを効率的に探索できるモデルを提案する。戦略をベクトル表現することで柔軟な探索を可能にし、進化的アルゴリズムの Quality Diversity アルゴリズムを用いることで効率的な探索を行う。提案するモデルによって、実世界のソーシャルネットワークの相互作用を正確に捉えることを目的とする。提案モデルは、既存手法と比較して実データとの整合性において優れた性能を示した。相互作用を行う際に多くの戦略を取るという結果から、効率的な戦略の多様性が示され、社会ネットワークの理解につながると考えられる。将来的には、より正確な予測モデルの設計からソーシャルメディア戦略の開発への情報提供まで、幅広い応用が期待される。

## 第2章 関連研究

### 2.1 エージェントベースのポリアの壺モデル

Ubaldi らのモデルは、社会的相互作用の複雑なダイナミクスを捉えるために、ポリアの壺モデルをベースとしている<sup>9)</sup>。このモデルは社会的相互作用における個人を「エージェント」、個人がそれぞれ持つ情報空間を「壺」として表現する。この壺には今後相互作用を行う可能性のあるエージェント、すでに相互作用したエージェントを示す「玉」が入っている。このモデルは次の3ステップで相互作用を行う。

**ステップ1** 発信側エージェントを選択する

**ステップ2** 着信側エージェントを選択する

**ステップ3** 発信側・着信側エージェント間で相互作用する

この3つのステップにより、相互作用のネットワークが生成される。これらのステップについて、以下に詳しく説明する。

#### ステップ1: 発信側エージェントを選択する

このステップでは、環境全体から発信側エージェントを一人選択する。壺に入っている玉の数を重みとして、ランダムにエージェントを選択する。そのため、壺が大きいほど、そのエージェントが選択される確率が高くなる。これは、優先的選択制を表現しており、コネクションの数が多いエージェントほど、新しいコネクションを構築する可能性が高くなることを表している。

#### ステップ2: 着信側エージェントを選択する

発信側エージェントが持っている壺の中から、ランダムに着信側エージェントを選択する。

#### ステップ3: 発信側・着信側エージェント間で相互作用する

発信側と着信側の相互作用によって、それぞれの壺の状態が変化する。過去に相互作用しているエージェントとの相互作用のしやすさを表す  $\rho$ 、新しいエージェントとの相互作用のしや

すさを表す $\nu$ および紹介するエージェントの選び方を表す戦略 $s$ の3つのパラメータが用いられる。Ubaldi のモデルの戦略には、(WSW)、Symmetric Sliding Window (SSW)、Asymmetric Sliding Window (ASW) などの6種類がある。3つのステップを踏んで相互作用の繰り返すことにより、ネットワークが成長する。その相互作用は図2.1の3つの状態のいずれかになり、それぞれ以下のように進行する。

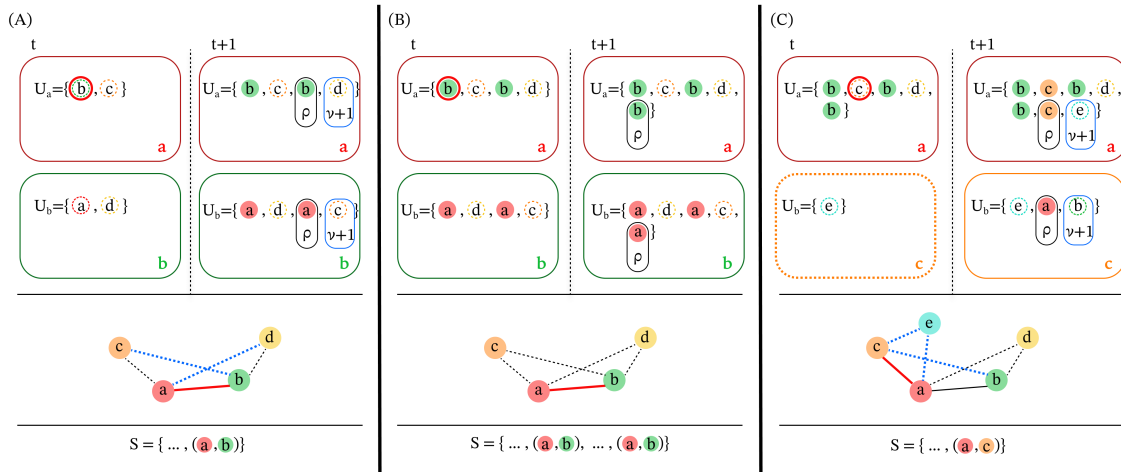


図 2.1: 相互作用の3つの状態

### 相互作用していない相手との相互作用 (A)

発信側エージェント  $a$  と着信側エージェント  $b$  が初めて相互作用する場合、両方のエージェントの壺にそれぞれ相手の玉を  $\rho$  個追加する。この強化ステップによって、既に相互作用した人と再度相互作用する確率が高くなる。初めての相互作用の場合のみ、両方のエージェントが持っている玉を  $\nu+1$  個ずつ交換する。この紹介ステップによって、両方のエージェントの隣接可能空間が広がり、過去に相互作用していない新しい人と相互作用する可能性がある。この交換方法は戦略  $s$  によって決められる。

### 既に相互作用した相手との相互作用 (B)

発信側エージェント  $a$  と着信側エージェント  $b$  が再度相互作用する場合も両方のエージェントの壺にそれぞれ相手の玉を  $\rho$  個追加する。初めて相互作用する相手でない場合は、互いの壺の中から交換する紹介ステップは行わないため、隣接可能空間は変化しない。



### 新しいエージェントの追加 (C)

どのエージェントとも繋がっていないエージェントが着信側として選ばれた場合、 $\nu + 1$  個のエージェントを新しく壺に追加する。その後は相互作用していない相手との相互作用 (A) の場合と同様に進行する。この場合は新しいエージェントが生成されるため、隣接可能空間をさらに広げることになる。

## 2.2 進化的アルゴリズム

進化的アルゴリズムは、生物の進化から着想を得た確率的探索手法である。問題に対して変化と選択に基づく世代交代を繰り返すことで、解の集団を進化させ最適解を得る。

### 2.2.1 遺伝的アルゴリズム

遺伝的アルゴリズムは、進化的アルゴリズムの中で最も一般的なアルゴリズムである。進化をモデルとした確率的探索手法の一つであり、問題の解を個体とする。個体ごとに遺伝子情報を与え、選択、交叉、突然変異といった操作を行うことで、個体の集団を進化させる。以下の流れで探索が行われる。

1. 初期集団を生成する
2. 適応度を評価する
3. 集団の中から親を選択する
4. 親の遺伝子を交叉させる
5. 子の遺伝子に突然変異を加える
6. 2~5 を繰り返す

### 2.2.2 Quality Diversity アルゴリズム

Quality Diversity(QD) アルゴリズムは、多様で高性能な解を探索するように設計されたアルゴリズムである<sup>14</sup>。このアルゴリズムでは、個体の振る舞いを特徴づけるベクトル Behavioral Descriptor(BD) を利用して、領域内の会に対する個体の新規性を定量化する。各領域はその領域に対して、可能な限り適合したエージェントになるように反復的に決定される。QD には二つの代表的なアルゴリズムがある。その一つは Novelty Search with Local Competition(NSLC)<sup>15</sup> である。このアルゴリズムは Novelty Search に局所的な競争の概念を導入している。個体をマッピングする領域は、BD 空間におけるユークリッド距離に基づき、非構造化アーカイブを用いて決定する。

二つ目の代表的なアルゴリズムに Muliti-dimensional Archive of Phenotypic Elites(MAP-Elites)<sup>16</sup>がある。このアルゴリズムでは、BD 空間はいくつかのグリッドに分けられる。各グリッドにはマッピングされた単一の最も適合したエージェントが記憶される。そして、BD 空間を埋めるように探索を進める。エージェントはグリッドから選択され、突然変異などのランダムな変化を受ける。そうしてできた新しい個体はその行動に対応するグリッドにマッピングされ、適合度が現在の占有者より高ければ記録される。

MAP-Elites はセル数が特徴空間の次元数に応じて、指数関数的に増加するため、高次元の BD 空間に対応するのは難しい。また、セル数が増えることによって、選択圧が低くなってしまうことも問題である。この問題を解決するために、MAP-Elites を高次元の特徴空間にスケールアップする CVT-Map Elites<sup>17</sup> が提案されている。このアルゴリズムは、BD 空間をグリッド上ではなく、重心ボロノイ分割によってボロノイ状のセルに分割する。この方法の場合、セル数を固定して BD 空間の次元を拡張することが可能になるため、高次元の BD 空間の探索において、MAP-Elites よりも優れた性能を示す。

## 第3章 手法

従来のモデルでは、人間がこの空間を探索する戦略をあらかじめ定義しておく必要がある。そのため、実際のデータに最も適合するパラメータの最適化探索は、これらの事前に定義した戦略に制約される<sup>8,9,13</sup>。この制限に対応するため、本研究では複数の戦略に対応し、これらを効率的に探索できるモデルを提案する。戦略をベクトル表現することで柔軟な探索を可能にし、進化的アルゴリズムの Quality Diversity アルゴリズムを用いることで効率的な探索を行う。提案するモデルによって、実世界のソーシャルネットワークの相互作用を正確に捉えることを目的とする。

既存のモデルでは、この問題を解決するために2つのことを行っている。(1) 戦略をベクトル表現することで、柔軟な探索を可能にし、探索空間を拡張する。(2) Quality Diversity アルゴリズムを用いることで、拡張された探索空間に適した効率的な探索を行う。

### 3.1 戦略のベクトル表現

Ubaldi らによって提案された既存モデルでは、戦略  $s$  は事前に定義された方法に制限されていた。これらの戦略は実世界の相互作用の一部を捉えることができますが、実際の相互作用に用いられる多様な戦略を全て表現することはできない。例えば、頻繁に相互作用するが、最近相互作用していないエージェントを優先する戦略は、実世界のネットワークを正確に表現するかもしれません。こように広い戦略を捉えるために、戦略を遺伝子を示すベクトルとして表現する。具体的には、戦略は最近度 (recency:  $r$ )、頻度 (frequency:  $f$ ) という二つの遺伝子として構成されるベクトルとして表され、それぞれの遺伝子は -1 から 1 の範囲内で値を取る。例えば、最近度が -0.5 で頻度が 0.5 の戦略ベクトルは、最近度が低い相互作用頻度が高いエージェントを優先する。図 3.1 に戦略ベクトルを使った相互作用の例を示す。図 3.1-(A) は最近度が高いが頻度が低い相互作用を優先する戦略の場合を表しており、図 3.1-(B) は最近度が低いが頻度が高い相互作用を優先する戦略の場合を表している。連続値で表現することで、探索空間のより包括的で柔軟な探索が可能になる。

### 3.2 進化的アルゴリズムによる探索

Ubaldi らのモデルは、あらかじめ決められた範囲内でパラメータ  $\rho, \nu, s$  の組み合わせの全探索を行なっている。最適なパラメータは、ネットワークの特徴を表す 10 個の指標を用いて

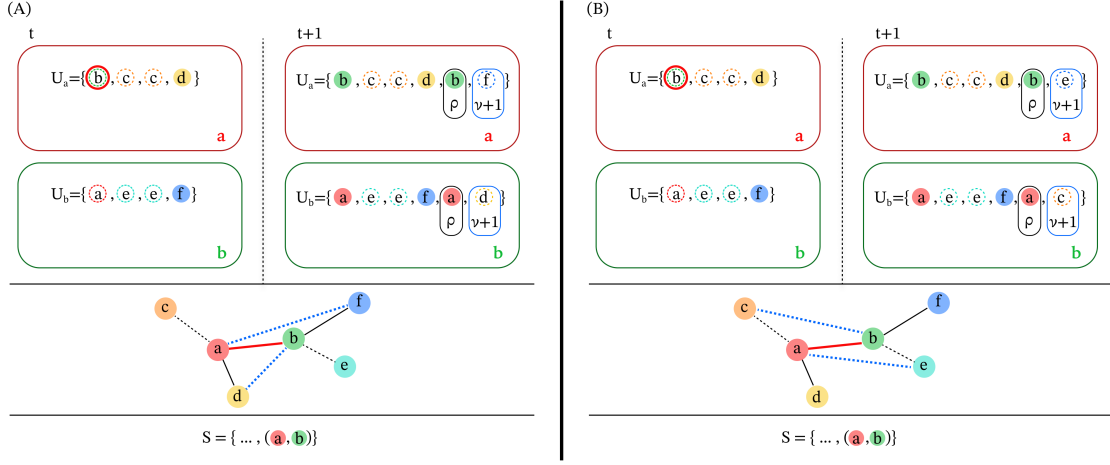


図 3.1: 戦略ベクトルを使った相互作用の例

評価される。モデルで生成したネットワークと実データのネットワークの指標の差を最小化するものを最適なパラメータとしている。

しかし、提案モデルでは戦略は連続値のベクトルとして表現されるため、潜在的な戦略の数が大幅に増加し、全探索するのは現実的ではない。この問題を解決するために、一つのパラメータセットを  $\rho, \nu$  と戦略を表す最近度  $r$ , 頻度  $f$  で構成される 4 次元ベクトルとして表し、進化的アルゴリズムを用いて実データに最も近いパラメータセットを効率的に探索する。

具達的には、遺伝的アルゴリズム (GA) と Quality Diversity (QD) アルゴリズムを用いる。遺伝的アルゴリズムでは、データとの距離  $d$  の逆数を適応度とし、選択にはルーレット選択を用いている。QD アルゴリズムとしては、高次元の BD 空間に対して MAP-Elites より高い性能を示す CVT-MAP-Elites を用いる。提案モデルによって作成されたネットワークを BD 空間にマッピングするために、graph2vec を用いて、ネットワークのグラフを 128 次元のベクトルに変換する。このベクトルを BD として用いる。

### 3.3 評価指標

モデルの目標データへの適応度を次式で計算される距離  $d$  を用いて評価する。

$$d = \sum_{p \in P} |P_{model} - P_{target}| \quad (3.1)$$

ここで、 $P = \gamma, C, Y, R, \langle h \rangle, G, OC, OO, NC, NO$  は、ネットワークの成長とダイナミクスを表現する 10 個の指標の集合である。モデルによって作成した全てのデータの中の距離  $d$  が最小のものが最適なパラメータとなる。

モデルの評価は、Ubaldi らが提案したモデルの生成するネットワークのトポロジーとそのダイナミクスに関する 6 つの指標<sup>9</sup> と、Monechi らが提案した 4 つの新規性に関する指標<sup>7</sup> を

用いて行う。Ubaldi らが提案している 8 つの指標の中で、詳細な記述と再現性が不足している指標を除いた 6 つの指標を用いている。

評価に用いた指標は以下の通りである。

- Heaps 則の指数  $\gamma$   
成長するネットワークで新しい要素が発見されるペースを定量化する
- クラスター係数  $C$   
ネットワークのクラスター性を表す
- 若さ係数  $Y$   
相互作用の新しさを表す
- 最近度  $R$   
相互作用の時間的な近さを表す
- 局所エントロピー  $\langle h \rangle$   
相互作用するエージェントの多様性を示す
- 擬似ジニ係数  $G$   
エージェントの注目度の格差を表す
- 4 つの相互作用の新規性と閉鎖性の指標  
 $\text{old-closed}(OC)$ 、 $\text{old-open}(OO)$ 、 $\text{new-closed}(NC)$ 、 $\text{new-open}(NO)$

相互作用の新規性は、相互作用が以前に発生しているかどうかによって定義され、old は複数回目の相互作用、new は初めての相互作用を示す。閉鎖性の概念は相互作用するエージェント同士が、隣接可能空間の中で共通のエージェントがいるかどうかに関連している。closed は共通のエージェントがいる場合、open は共通のエージェントがいない場合を示す。これらの指標をネットワークダイナミクス、構造、および新規性を定量的に評価するために用いる。

## 第4章 実験

本研究では3つの Research Question(RQ)を設定しており、その RQ に対応する実験を行なった。以下に RQ と実験の概要を示す。

**RQ1: 提案手法は Ubaldi らのモデルと比較して、モデルによって作成した合成データに対してどのような性能を示すか**

RQ1 に答えるためにランダムサーチ、遺伝的アルゴリズム、Quality Diversity アルゴリズムを用いた提案手法と全探索を用いる既存手法の4つの手法を比較した。ランダムサーチは、遺伝的アルゴリズムがランダムサーチよりも優れた利点を持つか確かめるために用いている。

**RQ2: 提案手法は Ubaldi らのモデルと比較して、実世界のデータの振る舞いをどの程度正確に再現できるか**

RQ2 に答えるために、提案手法と既存手法を実世界のデータに適用し、その結果を比較した。モデルによって生成した相互作用と3つの実データの相互作用との一致度を評価することによって、提案手法の有効性を評価した。RQ1 に対する実験と同様に、4つの手法を比較した。

**RQ3: 実世界のデータにおける隣接可能空間の探索では、どのような戦略が用いられているか**

RQ3 に答えるために、実データに対して、提案手法が見つけたパラメータから、隣接可能空間の探索に用いられている戦略を分析した。

### 4.1 データセット

本研究の実験には、モデルによって作成した合成データと実世界の振る舞いを示す3つの実データを用いた。

#### 4.1.1 合成データ

Ubaldi のモデルを用いて生成された合成データは、

Ubaldi のモデルでパラメータ  $(\rho, \nu) \in (5, 5), (5, 15), (20, 7)$ 、および  $s \in SSW, WSW$  を設定した。ここで、SSW は Symmetric Sliding Window 戦略を表し、壺の中に最も最近追加されたエージェントを優先する。一方、WSW は Weighted Sampling with Withdrawal 戦略を表し、壺の中からランダムにエージェントを選択する。そのため、確率的に頻度が高いエージェントが選択されやすい。これらのパラメータ  $\rho, \nu, s$  を組み合わせた合計  $6(= 3 \times 2)$  つのパラメータセットを用いて、それぞれ 20000 ステップの相互作用を行い、生成されたネットワークを合成データとする。この合成データを Ubaldi らの全探索と提案手法のフレームワークのランダムサーチ、遺伝的アルゴリズムおよび Quality Diversity アルゴリズムとの比較するための目標データとして用いた。

#### 4.1.2 実データ

実データとしては、X mention Network(TMN)、American Physical Society(APS) 共著ネットワークおよび MIXI Social Game Network(MIXI) の3つのデータセットを用いた。TMN と APS のデータセットは、Ubaldi らの研究<sup>9</sup>で使用され、公開されている。TMN と APS データセットの両方から、20000 の caller-callee 相互作用データを評価に使用する。MIXI データセットは、日本の主要なゲーム会社である株式会社 MIXI によって、提供していただいたオンラインソーシャルゲーム「モンスターストライク」のデータを用いている。MIXI データセットから 49,346,896 の caller-callee 相互作用データを評価に使用する。

### 4.2 パラメータ

合成データと実データの両方において、最適なパラメータの探索を行った。各手法に対して、指定した範囲内のアルゴリズムのハイパーパラメータとパラメータ  $(\rho, \nu, s)$  および  $(\rho, \nu, r, f)$  を探索した。

既存モデルでは、 $1 \leq \rho, \nu \leq 20$  および  $s \in \{SSW, WSW\}$  の 800 組のパラメータを全探索している。それぞれのパラメータセットで 20,000 ステップの相互作用を実行している。目標データとの距離  $d$  を最小に抑えたパラメータセット、すなわち目標データに最も近いネットワークを生成したパラメータセットを探索する。

提案モデルでは、 $1 \leq \rho, \nu \leq 30$  および  $-1.0 \leq r, f \leq 1.0$  の範囲で、ランダムサーチ、遺伝的アルゴリズムおよび QD アルゴリズムによりパラメータを探索する。ランダム探索では、合成データに対して 100 回、実データに対して 500 回の探索を行う。遺伝的アルゴリズム、QD アルゴリズムでは、合成データに対して 100 世代、実データに対して 500 世代に固定して探索を行う。遺伝的アルゴリズムでは、交叉率、突然変異率、個体数のハイパーパラメータをグリッドサーチにより最適化する。ハイパーパラメータは、交叉率  $C \in \{0.80, 0.85, 0.90, 0.95\}$ 、

突然変異率  $M \in \{0.01, 0.02, 0.03, 0.04, 0.05\}$ 、個体数  $N \in \{10, 20, 30, 40, 50\}$  の組み合わせの中から最適なものを選択した。

QD アルゴリズムでは、Python ライブラリの pyribs<sup>?</sup> を用いる。生成されたネットワークのグラフは、graph2vec アルゴリズムを用いて、ベクトルに変換する。CVT-MAP-Elites のセル数とグラフを変換したベクトルの次元数をグリッドサーチにより最適化する。セル数は 250, 500, 750、ベクトルの次元数は 64, 128, 256 の組み合わせの中から最適なものを選択した。変異率は CVT-MAP-Elites では初期値が設定されるが、CMA-ES アルゴリズム によって決められ、一定ではないため、ハイパーパラメータチューニングを行っていない。グリッドサーチによって、見つけた最適なハイパーパラメータを表 4.1 に示す。このパラメータを用いて、提案手法の探索を行った。

表 4.1: 最適なハイパーパラメータ

(A) 合成データ

	GA			QD	
	交叉率	突然変異率	個体数	セル数	BD の次元数
$(\rho = 5, \nu = 5, s = \text{SSW})$	0.8	0.01	30	500	128
$(\rho = 5, \nu = 5, s = \text{WSW})$	0.8	0.02	30	250	256
$(\rho = 5, \nu = 15, s = \text{SSW})$	0.95	0.03	20	500	128
$(\rho = 5, \nu = 15, s = \text{WSW})$	0.8	0.01	10	250	128
$(\rho = 20, \nu = 7, s = \text{SSW})$	0.8	0.04	50	750	64
$(\rho = 20, \nu = 7, s = \text{WSW})$	0.8	0.01	10	250	128

(B) 実データ

	GA			QD	
	交叉率	突然変異率	個体数	セル数	BD の次元数
TMN	0.85	0.05	20	250	64
APS	0.85	0.01	40	250	256
MIXI	0.85	0.04	30	250	128

## 4.3 実験結果と考察

### 4.3.1 RQ1: 合成データによる精度比較

6 つの合成データのそれぞれに対して、既存モデルの全探索、提案モデルのランダムサーチ、遺伝的アルゴリズム、QD アルゴリズムを用いて、最適なパラメータを探索した。この実験では、既知の正しいパラメータセットを用いた合成データ上でのパラメータフィッティングによって精度を評価する。6 つの合成データに対して実験を行ったが、全てのパラメー



タセットで同様の結果が得られたため、QD アルゴリズムの距離が最小のパラメータセット  $(\rho, \nu, s) = (5, 15, SSW)$  に対する結果のみを示す。

図 4.1 は、探索されたすべてのパラメータの距離  $d$  の最小値, 最大値, 四分位数を箱ひげ図で示している。

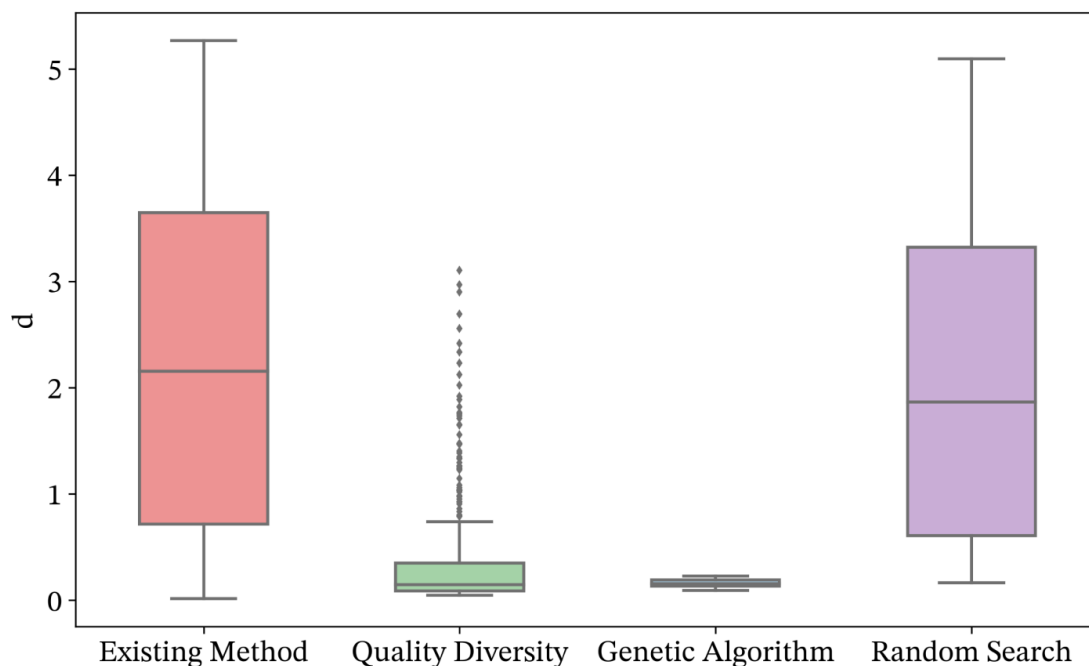


図 4.1: パラメータ  $(\rho, \nu, s) = (5, 15, SSW)$  の合成データと各手法の距離の箱ひげ図の比較

図 4.1 から、最小値はどのアルゴリズムでも大きな差はないことがわかる。また、距離の分散は既存手法と提案手法のランダムサーチが大きい。これらの手法は包括的な探索を行うため、距離の分散が大きくなる。次に QD アルゴリズム、遺伝的アルゴリズムの順に距離の分散が小さく、全体として距離が小さい。QD アルゴリズムと遺伝的アルゴリズムによって、高性能な解が探索されていることがわかる。QD は高性能かつ多様な解が見つまっているため、遺伝的アルゴリズムより距離の分散が大きくなっていると考えられる。QD で見つかった多様な解については、後述する RQ3 で詳しく分析する。

図 4.2 は、遺伝的アルゴリズムと QD アルゴリズムの 100 世代にわたる合成データ  $(\rho, \nu, s) = (5, 15, SSW)$  をターゲットとしたときの距離の変化を示す。波線は各世代において集団の中で最も最小の距離を示し、実線と薄い領域は集団全体の平均と第一四分位数と第三四分位数を表している。

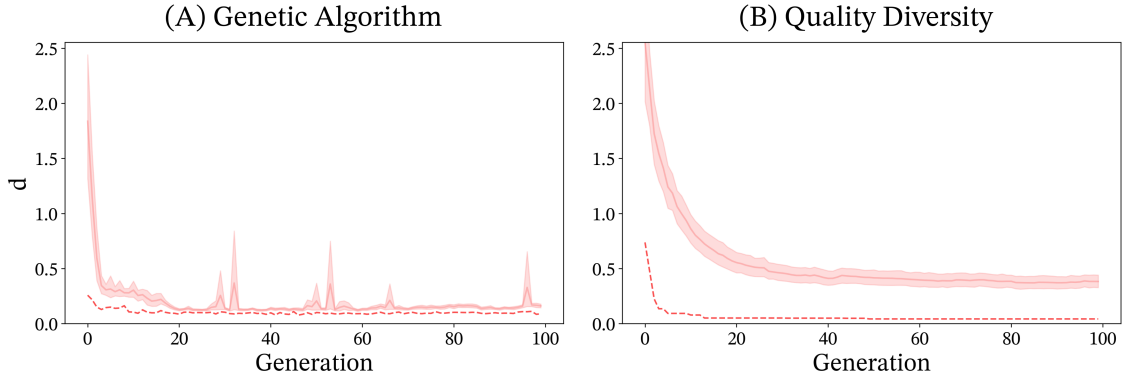


図 4.2: パラメータ  $(\rho, \nu, s) = (5, 15, SSW)$  の合成データに対する GA と QD の探索の時間経過

図 4.2 より、どちらのアルゴリズムでも距離の値が世代ごとに減少し、学習が進んでいることがわかる。しかし、QD アルゴリズムの方が遺伝的アルゴリズムより収束が遅く、集団全体の平均の距離は大きい。これは、QD アルゴリズムがより多様な解を保つように探索するためだと考えられる。また、遺伝的アルゴリズムの結果において、時折見られる平均の距離が前の世代より大幅に増える現象は突然変異によるものである。

表 4.2 は、それぞれの探索手法によって、見つけた最適なパラメータセットでの実行結果を示している。壺モデルの実行によって得られるネットワークにはランダム性があり、同じパラメータセットで実行したとしても、生成されるネットワークは全く同じにはならない。そのため、壺モデルを 10 回実行し、その結果の平均と標準偏差を示した。

表 4.2: Best fit values for synthetic data

Existing	QD	GA	Random
<b>0.068±0.026</b>	0.209±0.059	0.198±0.045	0.161±0.025

既存の手法による探索では、ターゲットデータとして用いる合成データのパラメータも探索されるため、最も高い適合度を示している。既存手法が示す距離の値は、壺モデル自体のランダム性によるものである。提案モデルでは、ランダムサーチが最良の結果であり、遺伝的アルゴリズムと QD もほぼ同等の結果を示した。QD アルゴリズムや遺伝的アルゴリズムの方がランダムサーチより効率的な探索が可能だが、この実験においてはわずか 4 次元のパラメータ空間を探索しているため、大きな差が出なかったと考えられる。

図 4.3 に評価に用いる 10 個の指標の結果を示す。QD アルゴリズムで距離が最小のパラメータセット  $(\rho, \nu, s) = (5, 15, SSW)$  と距離が最大の最大のパラメータセット  $(\rho, \nu, s) = (20, 7, SSW)$  の結果を示している。

最良のケースでは既存手法の方が距離が小さかったが、評価指標を見ると、ごく僅か差で

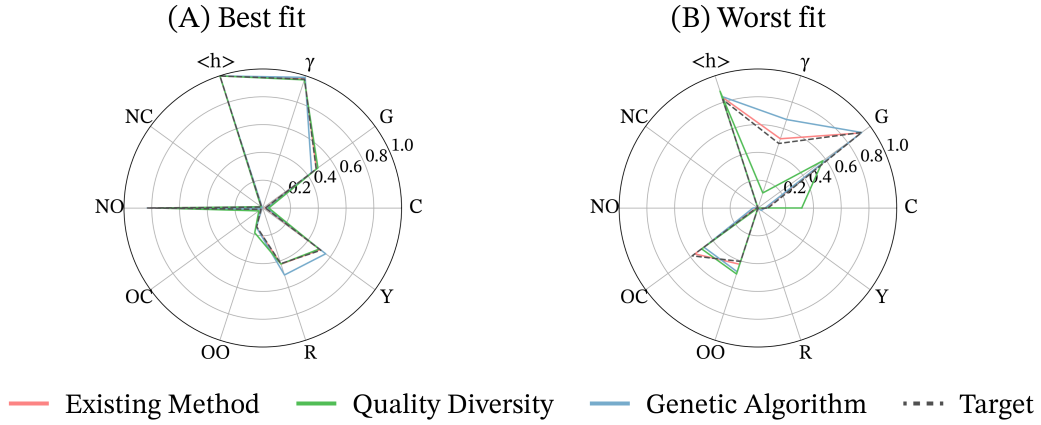


図 4.3: パラメータ  $(\rho, \nu, s) = (5, 15, SSW)$  の合成データに対する各測定指標の既存モデルと提案モデルの比較

あることがわかる。また、最悪のケースでも目標データにある程度フィットしている。これらの結果から、提案手法は目標データの振る舞いを再現できている。

#### 4.3.2 RQ2: 実データによる精度比較

最も実データに適合する 10 個の指標の結果を図 4.6 に示す。

表 4.3: Best fit values for TMN, APS and MIXI

	Existing	QD	GA	Random
TMN	$0.837 \pm 0.025$	<b><math>0.763 \pm 0.056</math></b>	$0.790 \pm 0.042$	$0.816 \pm 0.057$
APS	$0.603 \pm 0.037$	$0.520 \pm 0.028$	$0.534 \pm 0.033$	<b><math>0.518 \pm 0.024</math></b>
MIXI	<b><math>1.135 \pm 0.123</math></b>	$1.178 \pm 0.081$	$1.209 \pm 0.068$	$1.213 \pm 0.069$

#### 4.3.3 RQ3: 実データにおける隣接可能空間の探索に用いられている戦略

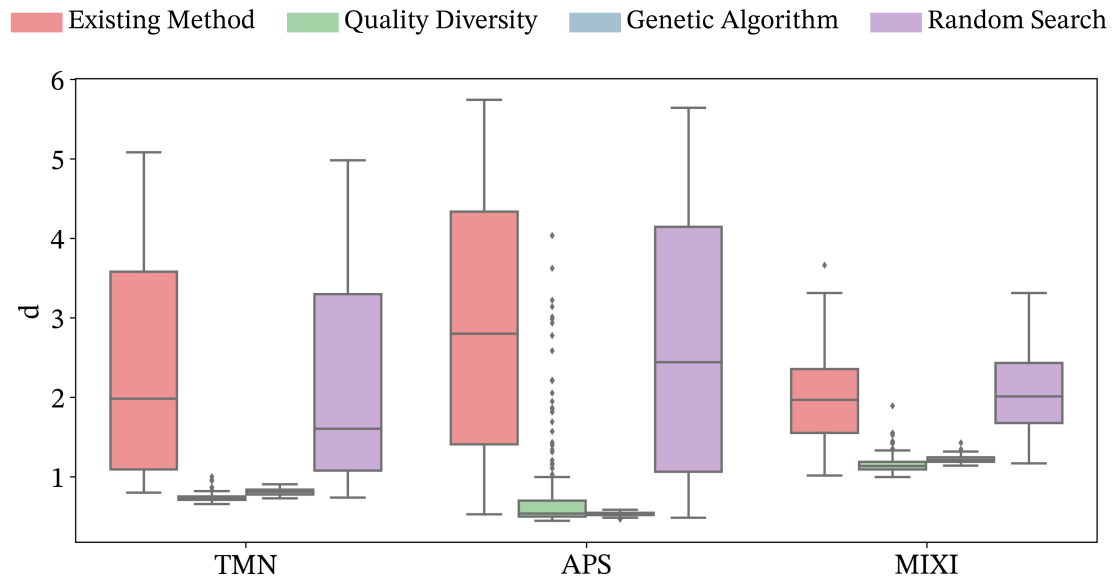


図 4.4: 実データと各手法の距離の箱ひげ図の比較

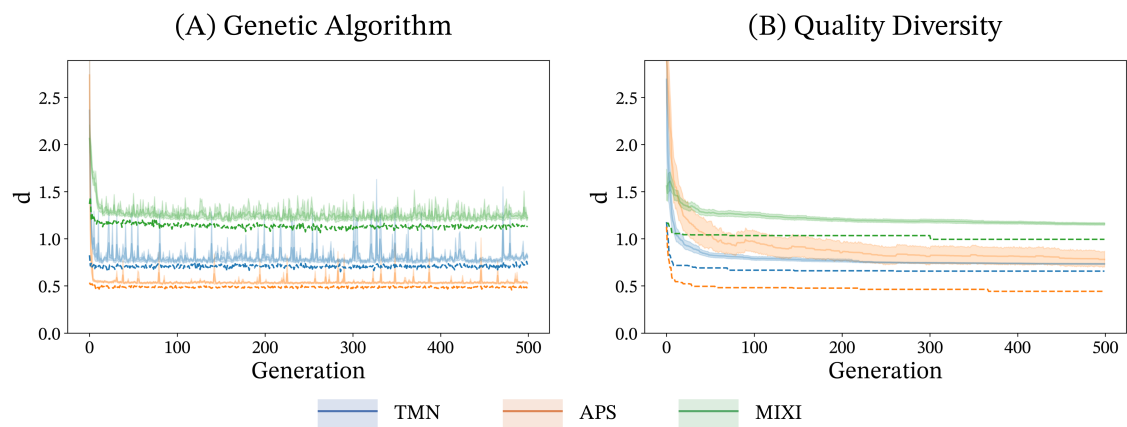


図 4.5: 実データに対する GA と QD の探索の時間経過

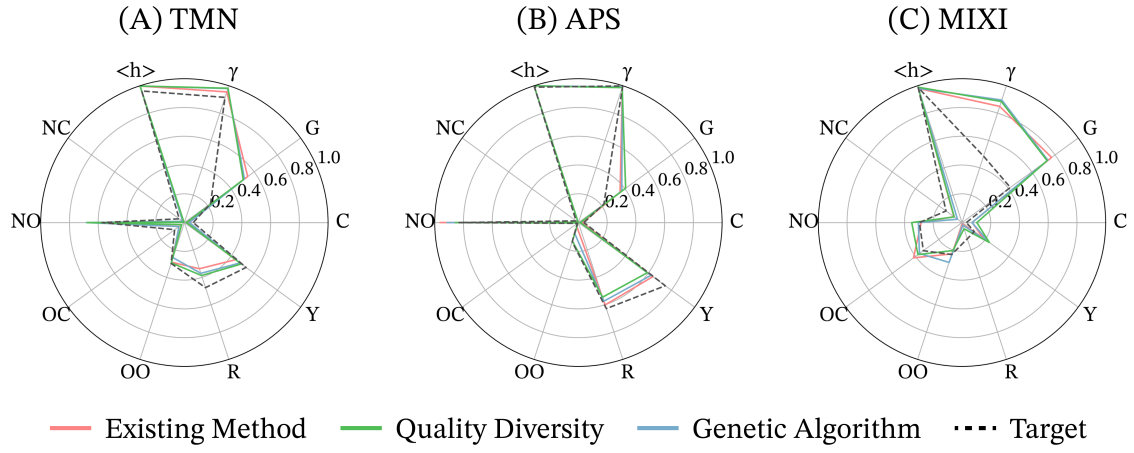


図 4.6: 実データに対する各測定指標の既存モデルと提案モデルの比較

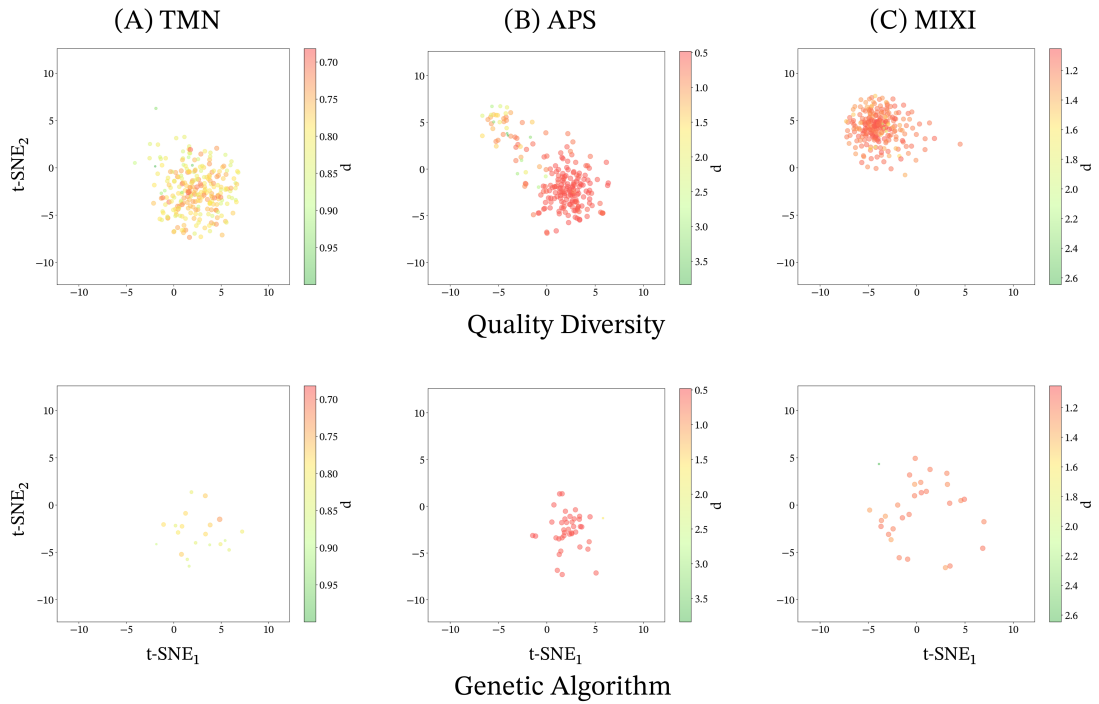


図 4.7: ベクトル化した QD と GA アルゴリズムによる探索で見つけたネットワークの可視化

表 4.4: Top-five best-fit parameters for each target dataset.

(A) TMN						(B) APS						(C) MIXI					
#	d	$\rho$	$\nu$	$r$	$f$	#	d	$\rho$	$\nu$	$r$	$f$	#	d	$\rho$	$\nu$	$r$	$f$
1	0.656	16.79	25.30	-0.12	-0.77	1	0.442	9.55	28.59	-0.47	-0.79	1	0.994	2.90	2.05	-0.91	0.34
2	0.660	17.20	26.04	0.28	-0.89	2	0.462	6.80	24.66	-0.59	0.13	2	1.024	2.56	2.67	-0.44	-0.01
3	0.661	19.34	29.53	-0.10	-0.33	3	0.463	7.14	23.62	-0.39	-0.09	3	1.030	2.42	2.87	-0.74	0.62
4	0.664	15.73	23.23	-0.65	0.02	4	0.464	7.84	26.96	-0.24	0.20	4	1.033	2.15	2.01	-0.92	-0.08
5	0.666	16.73	26.70	-0.54	-0.26	5	0.466	9.75	28.77	-0.60	-0.58	5	1.034	2.46	2.40	-0.93	-0.82

## 第5章 結論

## 謝辞

本研究の遂行にあたり、研究の着想から論文の完成に至るまで、熱心なご指導と多大なご助言を賜りました岡瑞起教授に深く感謝申し上げます。また、毎週のゼミや研究室での議論を通じて、様々なご助言をいただいた研究室の皆様にご心よりお礼申し上げます。特に、昨年卒業された須田幹大さんには、研究の方向性や壺モデルの実装に関する多くのご助言をいただきました。須田さんの経験と知識から多くのことを学び、研究を進める上で大変お世話になりました。また、岡部純弥さんには実験に関するサポートとご助言をいただきました。さらに、株式会社 MIXI の皆様には、データセットの提供していただきました。このように、多くの方々のご協力に恵まれ、本研究を遂行することができました。心より感謝申し上げます。



## 参考文献

- [1] Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. Science, Vol. 286, No. 5439, pp. 509–512, 1999. [eprint: cond-mat/9910332](#).
- [2] Albert-László Barabási. The origin of bursts and heavy tails in human dynamics. Nature, Vol. 435, No. 7039, pp. 207–211, 2005. [eprint: cond-mat/0505371](#).
- [3] Albert-László Barabási. Scale-Free Networks: A Decade and Beyond. Science, Vol. 325, No. 5939, pp. 412–413, 2009.
- [4] Réka Albert and Albert-László Barabási. Topology of evolving networks: local events and universality. Physical review letters, Vol. 85, No. 24, p. 5234, 2000. Publisher: APS.
- [5] Ginestra Bianconi, Richard K Darst, Jacopo Iacovacci, and Santo Fortunato. Triadic closure as a basic generating mechanism of communities in complex networks. Physical Review E, Vol. 90, No. 4, p. 042806, 2014. Publisher: APS.
- [6] Alexandru Topirceanu, Mihai Udrescu, and Radu Marculescu. Weighted Betweenness Preferential Attachment: A New Mechanism Explaining Social Network Formation and Evolution. Scientific Reports, Vol. 8, No. 1, p. 10871, July 2018.
- [7] Bernardo Monechi, Álvaro Ruiz-Serrano, Francesca Tria, and Vittorio Loreto. Waves of novelties in the expansion into the adjacent possible. PLOS ONE, Vol. 12, No. 6, p. e0179303, June 2017. Publisher: Public Library of Science.
- [8] Mikihiro Suda, Takumi Saito, and Mizuki Oka. Exploration and exploitation of the adjacent possible space for open-endedness. The 2022 Conference on Artificial Life, 2022.
- [9] Enrico Ubaldi, Raffaella Burioni, Vittorio Loreto, and Francesca Tria. Emergence and evolution of social networks through exploration of the Adjacent Possible space. Communications Physics, Vol. 4, No. 1, p. 28, February 2021.
- [10] Stuart A Kauffman. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA, 1993.
- [11] John Maynard Smith. Natural selection and the concept of a protein space. Nature, Vol. 225, No. 5232, pp. 563–564, 1970. Publisher: Nature Publishing Group UK London.

- [12] F. Tria, V. Loreto, V. D. P. Servedio, and S. H. Strogatz. The dynamics of correlated novelties. Scientific Reports, Vol. 4, No. 1, p. 5890, July 2014.
- [13] Mikihiro Suda, Takumi Saito, and Mizuki Oka. Agent-based model using adjacent possible space for generating waves of novelty. pp. 1251–1258. IEEE, 2022.
- [14] Justin K. Pugh, Lisa B. Soros, and Kenneth O. Stanley. Quality Diversity: A New Frontier for Evolutionary Computation. Frontiers in Robotics and AI, Vol. 3, p. 40, 2016.
- [15] Joel Lehman and Kenneth O. Stanley. Evolving a diversity of virtual creatures through novelty search and local competition. Proceedings of the 13th annual conference on Genetic and evolutionary computation - GECCO '11, pp. 211–218, 2011.
- [16] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. 2015. arXiv: 1504.04909.
- [17] Vassilis Vassiliades, Konstantinos Chatzilygeroudis, and Jean-Baptiste Mouret. Using Centroidal Voronoi Tessellations to Scale Up the Multi-dimensional Archive of Phenotypic Elites Algorithm. 2017. arXiv: 1610.05729.