

博士論文

地域クラスターのダイナミクスと
希少なつながりの役割に関する研究

山野 泰子

地域クラスターのダイナミクスと 稀少なつながりの役割に関する研究

概要

予期せぬ自然災害や金融危機など、変動が激しくなった現代市場において、企業間の相互作用に注目しつつ、変動に対する企業や地域クラスターの適応力を評価することに対する社会的関心が高まっている。本論文では、そうした社会的な要請を踏まえ、ネットワークを構成する要素間の動的推移と構造特性を捉える2つのノード指標を提案し、さらに、それらを企業間取引ネットワークにおける企業単体および密な取引関係を持つ企業群としてのクラスターの評価に実際に適用する枠組みを構築することで、新たな知見を抽出した。

第一章では、提案指標と密接に関連する社会背景や政策的課題として、日本の経済政策のなかで用いられてきた新陳代謝の概念と地域クラスターの定量的評価の必要性を述べるとともに、異質な分野間の連携と産業クラスターの構造転換が近年進んでいること、その背景にある企業間の遠距離交流が活発化していること等について論じた。また、こうした構造変化を捉える上でネットワーク科学の既存の評価指標とその限界を説明し、新たな指標等の開発による本論文の貢献と企業間の取引に関する動的推移解析、静的構造解析、クラスター単位の継承度や群間変動・進化から成る枠組みについて述べている。

第二章では、地域クラスターと企業間ネットワークに関する技術経営分野の先行研究を整理するとともに、本論文における提案指標の着想の契機となった、極相林のエコシステムと新陳代謝に関する概念を説明している。さらに、本論文で扱うネットワーク科学の分析手法として動的推移解析と静的構造解析、コミュニティ進化イベントの検出手法と提案指標を用いたイベントの予測の枠組みについて説明し、先行研究と対比しつつ、本論文の位置付けについて示した。

第三章では、ネットワーク中のミクロな動的推移を解析する独自の指標を提案している。具体的には、企業間取引関係の変動に基づいて定義した企業の新陳代謝度を定義し、それを東北地方、中部地方、九州地方の企業間取引の実データに適用して、企業間の相互関係のダイナミクスを定量的に解析した結果について述べる。企業の新陳代謝度と存続年数については、新陳代謝度の分位点別企業群間で多重比較検定を行った結果、新陳代謝度が0.1程度である企業群の平均存続年数が、統計的に有意に長いという知見が得られた。また、取引関係が固定的な企業および取引関係の変動が極端に大きい企業は、いずれもビジネス市場での存続が困難になることを明らかにした。

第四章では、ネットワークの静的構造に基づいた新たなノード指標を提案している。既存指標では、クラスター間をつなぐノードの検出は可能であるが、当該ノードがつかないでいるクラスター間の距離、すなわち、クラスター間のエッジ数や共通隣接ノード数が考慮されていない。本論文では、異質性の高いクラスターをつなぐことでネットワーク全体に多様性をもたらす企業を特定するため、距離が遠いクラスター間をつなぐノードを検出する指標PWを提案し、条件を変えて生成した複数のベンチマークネットワークを用いて、提案指標の妥当性を検討した。例えば、ネットワーク中のノードへの攻撃耐性に関する検証実験では、ネットワークの構造的な特徴を捉える8つノード指標の中で、提案指標PWが、稀少かつ重要なエッジを持つノードの検出において最も優れていることが明らかになった。

第五章では、第三章および第四章で提案した2つのノード指標を、東北地方の実際の企業間取引ネットワークにおける、クラスターの進化イベントの分析に適用して実験を行った。動的ネットワークにおいて、ミクロなノードのポジション価値が、マクロなクラスターの動態にどのように影響するかを明らかにすることは、これまでの手法では困難であった。本論文では、実験を通して、提案指標PWがクラスター進化イベントの予測に最も大きく寄与していることを明らかにした。また、PW値の高いブローカーと名付けた企業には、隣接ノードが少ない、移行性が高い、移動性が高いという3つの特徴があることを特定した。さらに、クラスター進化の系譜を表す独自のクラスター進化マップを用いることで、3つの典型的なクラスターパスの抽出に成功した。その上で、新陳代謝度、Z値、P値、PW値、売上高、産業エントロピー、地域エントロピー、クラスターサイズの8つの指標の推移を示すレーダーチャート、およびクラスター進化イベントの出現傾向を用いることで、それぞれのクラスターパスが持つ特徴を捉えられることを示した。

第六章では、提案指標や枠組みを用いた解析結果を踏まえ、結論として、適切な進化を遂げているネットワークにおいては、安定と変化という2つの拮抗する力のバランスが保たれていること、提案指標PWを用いて検出されるブローカー企業が地域の産業構造を進化させる触媒として重要な役割を果たしていること等を述べている。次に、本論文を踏まえた政策提案として、提案指標を用いた定量的アプローチによる政策立案や運用プロセスへの導入、クラスターパスの特性に応じた施策の投入による政策効果の引き上げ、提案指標を用いた多様性をネットワークに取り込むことに貢献するブローカー企業群の特定と条件検討、該当企業との政策対話等をまとめている。

目次

第1章	序論	1
1.1	背景・課題	1
1.1.1	ビジネス生態系における企業の新陳代謝	1
1.1.2	産業クラスターの構造転換と地域創生の課題	3
1.1.3	データサイエンスの導入によるネットワークの評価	4
1.2	着想の原点と本論文のアプローチ	8
1.3	貢献	10
1.4	本論文の構成	10
第2章	関連研究および手法の着想	13
2.1	地域クラスターと企業ネットワーク研究	13
2.1.1	地域クラスター戦略：産業集積の利点と課題	13
2.1.2	地域ネットワーク研究：つながりの価値とブローカーの役割	15
2.1.3	地域クラスターと企業間相互作用のダイナミクス	19
2.2	生物と経済の接合過程に見る持続的社會モデルへの試案	22
2.2.1	生物界の新陳代謝	22
2.2.2	産業界の新陳代謝	23
2.3	ネットワーク解析の分析方法	26
2.3.1	動的推移解析：環境適応と影響予測	26
2.3.2	静的構造解析：ノード指標とコミュニティ構造	28
2.3.3	コミュニティ進化の動的モデル	31
2.4	本論文の位置づけ	33
第3章	提案指標1：新陳代謝度（リンク推移を反映したノード指標）	35
3.1	提案指標の定義と分析手法	35
3.1.1	企業間取引に基づく新陳代謝の基本概念	35
3.1.2	群間の統計的有意差検定	37
3.1.3	非負値テンソル因子分解のアルゴリズム	39
3.2	検証実験1：企業間取引の新陳代謝度の特徴	41
3.2.1	データセット	41
3.2.2	企業の新陳代謝度と存続年数との相関	42
3.2.3	新陳代謝度と地域経済	51

3.3	検証実験 2：新陳代謝度の有意差検定と取引パターン分析	54
3.3.1	データセット	54
3.3.2	存続年数別新陳代謝度の分布	55
3.3.3	新陳代謝度の統計的有意差検定	57
3.3.4	非負値テンソル因子分解による取引変動のパターン分析	62
3.4	考察と結論	64
第 4 章	提案指標 2：コミュニティ異質性を反映したノード指標 PW	67
4.1	比較指標の定義と本章の位置付け	67
4.1.1	コミュニティ内外のリンクとノードの役割	68
4.1.2	コミュニティの構造情報を用いたリンク予測	68
4.1.3	本章の位置づけ	70
4.2	提案指標の定義と分析手法	70
4.2.1	コミュニティ異質性を反映したノード指標	70
4.2.2	隣接行列の更新による計算行程	71
4.2.3	P 値と PW 値のノード識別力の比較	73
4.3	検証実験	74
4.3.1	データセット	74
4.3.2	実験内容と結果	75
4.3.3	ノード指標別順位とネットワーク可視化	76
4.3.4	ネットワークへの攻撃耐性とノードの順位相関	80
4.4	考察と結論	88
第 5 章	実データの適用によるクラスター進化とブローカーの役割の特定	89
5.1	指標の定義と分析手法	89
5.1.1	クラスター進化に関する分析指標の定義	89
5.1.2	クラスター継承度の回帰分析	93
5.2	検証実験	97
5.2.1	データセット	97
5.2.2	ブローカーの特徴	101
5.2.3	構造的ノード指標を用いたクラスター進化マップ	104
5.3	クラスターパスとブローカーのケーススタディ	109
5.3.1	クラスターパスの安定性	109
5.3.2	地域多様性を持つ定常クラスター	110

5.3.3	ローカルハブを中心とした地域クラスター	111
5.3.4	ブローカーが局在する小規模アジャイルクラスター	113
5.3.5	Role 分布のクラスターパス別年次推移	114
5.3.6	クラスターパス毎に異なるブローカーの役割	116
5.3.7	ブローカーの維持年数と復興需要	116
5.4	クラスター継承度の予測	119
5.4.1	探索的データ解析	119
5.4.2	統計的回帰分析によるクラスター継承度の予測	121
5.4.3	機械学習モデルによるクラスター継承度の予測	123
5.5	考察と結論	124
第 6 章	考察	127
6.1	変動する市場環境への適応力に関する考察	127
6.1.1	企業取引の新陳代謝度の企業パフォーマンスへの影響	127
6.1.2	拮抗する二つの力	127
6.2	異質性を取り入れる仕組みとイノベーションの可能性	128
6.2.1	ブローカーの検出手法	128
6.2.2	ノードのポジション価値とクラスター進化	128
6.3	提案手法の組み合わせによる新たな示唆	129
6.3.1	越境型クラスターと多様性維持の仕組み	129
6.3.2	政策的インプリケーション	130
6.4	手法の限界と今後の研究課題	133
6.4.1	データおよび指標設計上の限界	133
6.4.2	提案手法の適用対象と影響伝播モデルへの応用	134
第 7 章	結論と展望	135
7.1	本論文の結論	135
7.2	今後の展望	135
	謝辞	137
	参考文献	139
	付録 A	155
A.1	第 3 章の補足資料	155
A.2	第 4 章の補足資料	168

A.3 第5章の補足資料	173
A.4 ソースコード	182

目次

1.1	構造的空隙を識別する 3 指標の比較	6
1.2	facebook 上の 10 人のユーザーの友人ネットワークと共通の友人	7
2.1	極相の古典的な遷移, ギャップダイナミクス, 非均衡理論の概念図 [1]	23
2.2	コミュニティ内外のリンク度を測定するノード指標 Z 値と P 値	30
2.3	コミュニティ差異を反映したノード指標 PW の概念図	31
3.1	企業間取引に基づく新陳代謝の概念図	36
3.2	NTF の概念図	39
3.3	新陳代謝度の分布	43
3.4	新陳代謝度と存続年数の分布の地域比較	46
3.5	Changer と Holder の存続年数と新陳代謝度	50
3.6	Changer と Holder の産業別取引先企業数分布	50
3.7	新陳代謝度と売上高の県別比較	52
3.8	参入・退出取引の産業別年次推移	53
3.9	産業別発注社企業	54
3.10	東北地方の新陳代謝度と存続年数の分布	56
3.11	東北地方の新陳代謝度と存続年数の分布 1	56
3.12	東北地方の新陳代謝度と存続年数の分布 2	56
3.13	新陳代謝度の分位点と存続年数の多重比較検定結果	60
3.14	新陳代謝度分位水準での存続年数の 95%信頼区間	61
3.15	NTF の近似誤差と繰り返し回数	63
3.16	NTF の近似誤差と基底数	63
3.17	NTF 因子分解結果	65
4.1	コミュニティ関連度の概念図	70
4.2	隣接行列の更新によるコミュニティ関連度取得	71
4.3	P 値と PW 値のノード識別力の比較	73
4.4	各ネットワークにおけるコミュニティの位置関係	77
4.5	RC ネットワークの指標別ランク	78
4.6	LFR ネットワークの指標別ランク	79
4.7	東北ネットワークの指標別ランク	80
4.8	ノード削除後の平均最短経路長による順位相関の指標別比較	82

4.9	最大次数の異なる LFR ネットワークでの順位相関	86
4.10	ノード攻撃後のネットワーク径の変化	87
5.1	2016 年の地域別・クラスター別企業数	100
5.2	2016 年の産業別・クラスター別企業割合	100
5.3	2016 年における東北地方の企業の指標別取引数	101
5.4	上位 100 企業の指標別年次推移	102
5.5	上位 200 企業の指標別クラスター内安定性	103
5.6	上位 200 企業の指標別被消失数	104
5.7	コネクターの分布とクラスターパス	106
5.8	ブローカーの分布とクラスターパス	106
5.9	ローカルハブの分布とクラスターパス	107
5.10	新陳代謝度が高い企業の分布と分岐を含むクラスターパス	108
5.11	クラスターパスの安定性	109
5.12	クラスターパス 1 の特徴推移とパス平均値の比較	110
5.13	クラスターパス 2 の特徴推移とパス平均値の比較	112
5.14	クラスターパス 3 の特徴推移とパス平均値の比較	113
5.15	Role 毎に見たクラスターパス別企業数の年次推移	115
5.16	クラスターパス毎のブローカーの維持年数	116
5.17	売上伸長率のクラスターパス別年次推移と分位点分布	118
5.18	各変数の値とデータ分布	120
5.19	各変数における外れ値の分布	120
5.20	変数間の相関関係	121
5.21	クラスター継承度に対する各説明変数の重要度	124
A.1	地方別新陳代謝度の分布	155
A.2	新陳代謝度の分位点水準での存続年数の分布と TK 検定 (全企業)	157
A.3	新陳代謝度の分位点水準での存続年数の 95%信頼区間 (全企業)	158
A.4	存続年数別企業群における新陳代謝度の 95%信頼区間	159
A.5	存続年数別企業群における新陳代謝度の有意差検定	160
A.6	新陳代謝度分位点における企業年齢の多重比較検定	160
A.7	新陳代謝度分位点における企業年齢の群間比較の 95%信頼区間	161
A.8	企業年齢分位点における新陳代謝度の多重比較検定	161
A.9	企業年齢分位点における新陳代謝度の群間比較の 95%信頼区間	162

A.10 東北地方の新陳代謝度 2 と存続年数の分布	164
A.11 新陳代謝度 2 の分位点と存続年数の多重比較検定結果	165
A.12 新陳代謝度 2 の分位水準での存続年数の 95%信頼区間	166
A.13 Facebook 10 ego users' network with 15 communities	168
A.14 Facebook three ego users' network with 12 communities	168
A.15 Comparison of the distributions of P and PW in Facebook networks	169
A.16 PW 値トップ 400 企業の分布とクラスターパス	173
A.17 各年における東北地方の企業の指標ランク別取引数	174
A.18 産業別・クラスター別の企業割合	175
A.19 地域別・クラスター別の企業数	176
A.20 クラスターパスの特徴推移（その他）	178

表 目 次

1.1	本論文の枠組み	11
3.1	データ概要	42
3.2	地方別企業の存続年数と固定化取引	44
3.3	新陳代謝度と企業の存続年数	45
3.4	地域・産業分類別, 新陳代謝度と存続年数	47
3.5	地域・県別, 新陳代謝度と存続年数	48
3.6	Changer と Holder の比較	49
3.7	Changer と Holder の産業別・県別新陳代謝度と存続年数	51
3.8	東北地方の企業間取引データ	55
3.9	独立 2 群間の有意差検定	57
3.10	新陳代謝度の分位点別企業数と基本統計量	58
3.11	新陳代謝度の分位点水準での多重比較検定	59
3.12	変数のカテゴリと分位点における閾値	62
4.1	RC ネットワークにおけるコミュニティペアの関連度	76
4.2	LFR ネットワークにおけるコミュニティペアの関連度	76
4.3	東北ネットワークにおけるコミュニティペアの関連度	76
4.4	コミュニティ数の異なる RC ネットワークでの順位相関	83
4.5	コミュニティサイズの異なる RC ネットワークでの順位相関	84
4.6	リンク率の異なる LFR ネットワークでの順位相関	85
5.1	クラスターパス一覧	90
5.2	クラスター継承度予測の機械学習モデルにおけるパラメーター範囲	96
5.3	東北ネットワークのデータ概要	98
5.4	地域別平均特徴量	99
5.5	産業別平均特徴量	99
5.6	2016 年の東北ネットワークにおけるクラスター内リンクのみを保有する企業数	105
5.7	クラスターパス 1 の産業別・地域別企業数	111
5.8	クラスターパス 2 の進化イベント	111
5.9	クラスターパス 2 の産業別・地域別企業数	112
5.10	クラスターパス 3 の産業別・地域別企業数	114
5.11	ブローカーの維持年数と平均評価値	117

5.12	適応条件を満たすブローカーの維持年数と平均評価値	118
5.13	クラスター継承度とノード評価指標の回帰分析結果	122
5.14	VIF と説明変数の選択	122
5.15	クラスター継承度とノード評価指標の重回帰分析結果	123
5.16	各アルゴリズムの統計量	123
A.1	独立 2 群間の有意差検定 (全企業)	156
A.2	分位点別企業数と平均新陳代謝度 (全企業)	156
A.3	新陳代謝度の分位点水準での存続年数の有意差検定	158
A.4	Pairwise comparisons using Games-Howell test	163
A.5	Pairwise comparisons using Dwass-Steele-Critchlow-Fligner all-pairs test	163
A.6	新陳代謝度 2 の分位点別企業数と基本統計量	165
A.7	新陳代謝度 2 の分位点水準での多重比較検定	167
A.8	Node values in the Facebook networks with P and PW	169
A.9	Node values in the RC networks with six indexes	170
A.10	Node values in the LFR networks with six indexes	171
A.11	Node values in the Tohoku networks with six indexes	172
A.12	クラスターペアの Jaccard 類似度と翌年のクラスター ID	177
A.13	クラスターパス 1~6 の特徴量	179
A.14	クラスターパス 4 の産業別・地域別企業数	180
A.15	クラスターパス 5 の産業別・地域別企業数	180
A.16	クラスターパス 6 の産業別・地域別企業数	180
A.17	クラスターパス 7 の産業別・地域別企業数	181
A.18	クラスターパス 8 の産業別・地域別企業数	181

第1章 序論

本章では，本論文の背景，目的，貢献及び構成について述べる．

1.1 背景・課題

1.1.1 ビジネス生態系における企業の新陳代謝

変動する市場環境にどう対応するかは企業にとって最も重要な課題の一つである．2011年の東北地方に甚大な被害をもたらした東日本大震災では，サプライチェーン崩壊の影響による日本企業へのショックが，被災地に留まらない日本全国規模にまで及んでいたことが可視化され，広く認識されるようになった [2]．これはサプライチェーンが広範囲で相互につながりあっており，危機の際に連鎖的反応を起こすこと，特定供給源への調達の集中など過度の効率化にはリスクが伴うことを示すものでもあった．新型コロナウイルスの感染拡大による活動制限の影響で，サプライチェーン寸断のリスクは一層深刻化し，2021年3月に行われた日本経済新聞による社長100人へのアンケート調査では，国内に工場を持つ企業の8割が供給網の分散化や柔軟化等の見直しに着手していることが判明している．高度につながりあった近年の市場経済における顕著な傾向は，特定地域における災禍や経済危機，貿易摩擦等の影響が，発生源となった地域を超えて大規模な範囲に及ぶことである．さらに，情報技術の発達により，世界中の情報がインターネット上を飛び交い，情報取得のコストが下がると共に変化のスピードは速くなった．情報はより遠くに，より速く広がるようになったといえる．大規模で予測不能な市場変動以外にも，日々の受発注やアライアンス，連携ないし協力といった企業間関係の中で相手企業からくる変動要因，あるいは直接取引関係がなくても競合関係にある企業の動向など，企業活動に影響を与える外部の変動要因は多い．本論文の第一の着眼点は，こうした変動の最中であって生き残るのは，相互に結びついたビジネス生態系とも言える企業間のネットワークにおいて，他企業との取引関係をどのように推移させる企業かということである．テクノロジーが進歩し変化が加速している現在の市場環境において，既存の取引関係に固定されている企業が，新しい取引先を開拓し供給網を再構築し続ける企業と比べて，存続年数を始めとする企業パフォーマンスにどのような違いがあるかは，日本経済の持続的な発展を考える上でも重要な視点と思われる．

Teeceは急速に変化する市場環境に対処するために，内部および外部の能力を統合し，構築し，再構築する企業の能力をダイナミック・ケイパビリティ (dynamic capability) と定義した [3]．環境の変動に対応する企業の能力に着目した同様の観点は，サプライチェーンの動態に応じた企業の柔軟性 (flexibility) や鋭敏性 (agility) [4]，吸収力 (absorptive capacity)

[5, 6], 回復力 (resilience) [7] をテーマにした研究にも見られる。しかし、これらの研究はアンケート調査やケーススタディに基づくものが大半であり、分析対象が限定されていること、定性的分析による理論構築に力点が置かれていることから、研究者によって概念の定義や適用範囲が異なり、統一の見解が得られにくいという課題がある。サプライチェーンに関する研究において、明確な定義に基づく測定可能な指標とバイアスの少ない網羅的データを用いた定量的、客観的な分析と評価が不足していることは、近年の文献レビューでも度々指摘されている [8, 9, 10]。本論文ではこうした課題を踏まえ、大規模なサプライチェーンの動態から企業を評価する新たな定量的指標として、新陳代謝度を提案する。

日本企業の取引関係に基づく定量的分析は、企業間取引データの整備が進み、企業情報提供サービスが拡充するにつれて広く行われてきた。2006年の中小企業白書には、産業のグローバル最適立地が進む中で、国内の取引構造の変化にも留意し、企業変革に取り入れていくことの重要性が述べられている [11]。その背景には、高度成長期以来、日本の特徴的な産業構造と言われてきた「親企業一下請企業」の長期安定的な取引構造が、製造拠点の海外進出・国際分業が進展する中で、分散化・多様化してきたという事実認識がある。「新陳代謝」は、近年の日本の経済政策の中で度々用いられてきた概念である。岡崎は日本経済が高い成長率を維持していた19世紀末以降の約100年間において、主導産業の交代をはじめとする大きな産業構造の変化が度々見られたことを明らかにし、日本の経済成長は同質的な企業や産業の規模拡大によるものではなく、産業の新陳代謝を伴うものであったと結論づけている [12, 13]。新陳代謝を促す経済政策は、東日本大震災の影響でサプライチェーンが寸断し、多くの企業が倒産・廃業を余儀なくされた2011年以降、起業・転業による経済の新陳代謝 [14]、開業率・廃業率に基づく産業の新陳代謝 [15, 16, 17]、経営者の世代交代、親族外への事業承継等による経営資源の企業内移転を表す企業内部の新陳代謝 [18, 17, 19] といった言葉で、中小企業白書の中でも頻繁に用いられている。その時々で定義や使われ方は異なるが、「新陳代謝」は、企業ないし産業の動的発展をもたらす現状打開策として着目され続けてきたことが伺える。

このように経済成長を促進する観点から変動する市場環境に対応する産業ないし企業のダイナミクスを重視する視点は、産業毎に異なる技術進歩の速度に着目しクロックスピードの概念を提唱したFineの視点 [20] や、新興企業が次々に現れ世代交代のスピードが早いディスク・ドライブ業界に着目して、破壊的イノベーションの概念を提唱したChristensenの視点 [21] にも通じるものがある。そこにあるのは、直線的経済成長を重視する経済成長の牽引力としてのプレイヤー交代、およびそれに伴う構造変化への着目である。そこでの交代ないし進歩のスピードは、速ければ速いほうが有利であることが前提となっている。

本論文では、「産業の新陳代謝」の概念を最初に提唱したAyresの着想に立ち返り、直線的成長モデルとは異なる視点からの分析を行う。Ayresは生物の有する生命維持機構としての新陳代謝と経済システムの制御機構のアナロジーからヒントを得て、この概念を提唱している [22]。そこには物質の流れを伴うエコシステムとしての経済圏、エネルギー効率化など持続可能な発展に向けた産業再編を重視する、有限なりソースを持続的に使う生存戦略としての視点がある。Ayresが重視したのは、交代を伴う系が定常的に維持される仕組みであり、その意味で生命体の新陳代謝の本質を捉えていたと言える。本論文で提案する新陳代謝度は

Ayres の視点に倣い、ビジネス生態系の中で持続可能な発展を遂げる、交代を伴う系としての企業の生存戦略、すなわち、個別の企業の取引関係のダイナミクスに着目していることに特徴がある。詳細は第3章で論ずるが、そこから得られた示唆は、取引先を交代するスピードには最適な値が存在するという、遅すぎてもいけないが速ければよいというものでもないということであった。

1.1.2 産業クラスターの構造転換と地域創生の課題

本論文の第二の着眼点は、イノベーションを誘発する異質な情報の流路はどのようなネットワーク構造によるものかということである。イノベーションの概念を提示した Schumpeter は、生産の場において企業者が遂行する「新結合」を経済発展の起動力とした [23, 24]。異質な存在とつながることができる環境がイノベーションに重要であることは、多くの先行研究によって示されているが、どのようなつながりが望ましい異質性であるかは時代に応じて変化している。

戦後の日本では、生産性向上を優先する産業政策が取られ、規模の経済性が追求されたため、地域は大手企業の量産工場の地方展開の場となり、地域固有の多様性や異質性は顧みられない状況が続いた [25, 26]。たとえば、東北地方の産業クラスターは 1972 年の東北自動車道開通に伴う、関東地方の電気・電子産業や機械産業関連の工場移転により成長したが、そのうちの多くは東京本社の統制下にある、研究開発部門を持たない支社であった。このため、当時の東北地方の産業集積は、地域のローカルネットワークよりは東京との垂直的つながりに従って作られたと言える。その後、地域の主要企業周辺に中小企業クラスターが成長し、東北地方の産業構造が形成された。高度経済成長期の日本の産業集積は、ものづくりに不可欠な要素技術のネットワーク化を容易にすることで、企業間の高度な擦り合せの実現に寄与していた。原材料・素材から部品・装置を経て最終製品に至る過程での企業間競争により、それまでの直列型の下請け構造が変化し、産業集積の内部において、新たな取引先開拓を含む取引関係の多面展開とネットワーク化が進んだことが、新技術開発の素地となっていたのである。このためバブル崩壊後の 1990 年代から 2000 年代初期にかけて、日本の競争力の源泉として政府が重視していたのは、産業集積の維持・強化であった [27]。

これに対し、経済産業省が 2001 年に施行した産業クラスター政策は、競争力のある地域産業・企業の発展支援を行うことで、産業集積の質を向上することを目的としている [28, 26]。産業クラスター政策は、クラスター内では知識が迅速かつ広範に流れ、非公式な対話も含めた交流や融合の機会が多くなることで柔軟な企業間の関係構築をしやすくなるとする Porter のクラスター戦略 [29, 30, 31, 32, 33] に基づいている。その中心思想は、企業等が集積する立地への着目というよりは、企業や産学官など様々なプレイヤー間の協力や競争関係のネットワークによって生み出される、イノベティブな事業環境構築にある。産業クラスター政策はネットワーク形成支援に重点を置いた政策であり、なかでも単なる受発注取引ではなく、異なる技術と技術、知識と知識の連携や技術移転を図ることを重視していると捉えられる [34, 35]。組織間の流動性の低さや、固定的なやりとりに閉じてしまいがちな社内ネットワーク、セクター間の縦割り構造の壁を乗り越えるには、さまざまな経済主体が集積するこ

とで、クラスター内の知識フローを増やすことが必要だと考えられてきたのである [29, 36].

しかし、近年における知識連携や技術移転をもたらすネットワーク形成は、地理的近接性が有利に働く特定分野のみで生じるとは限らない。Porter が分析した 1990~2000 年のクラスター、そしてそれをいち早く取り入れた欧米を始めとする先進国のクラスター政策が想定していたのは、特定技術を中心とする分野特定型のクラスターであった [32, 30, 37]. しかし、近年の先進国市場においては、既存技術の漸進的成長をもたらすクラスター内部のつながりの強化に加え、異質なクラスターをつなぐ新たな融合分野を生み出すことが一層重要になってきている [38]. たとえば、2020 年 10 月にドイツのメルセデスベンツがスイスのマイクロ・モビリティ・システムズ社と共同開発した電動キックボードの『e スクーター』はその一例と言える。高級車と電動キックボードでは取引先企業や顧客層が大きく異なるため、両社の取引ネットワークや所属クラスターは相違が大きいことが明らかである。しかし両社の融合分野においてマイクロモビリティの最新技術に耐久性とブランド力が加わり、ラストマイルにおける移手段の高付加価値化が実現した。日本においても、トヨタとソフトバンクの提携、日産と DeNA による自動運転の共同開発など、高度経済成長を牽引してきた大企業と新興企業との従来は考えられなかった異質な組み合わせでのビジネス連携が相次いでいる。石炭、鉄鋼、自動車といった旧来型製造業の衰退が進む米国ラストベルト（錆び付いた工業地帯）では、最先端のマテリアルサイエンス導入により、古い産業技術に新たな連携関係が生まれ、地域が新生しているとの報告もある [39]. そこでは古い工業地域の廃工場や安価な不動産が、大都市の知識労働者や起業家を誘引する魅力的な住居や仕事場に変貌しているという。こうした異質な分野間の連携は従来の産業から見ると、既存の交流関係とは異なる「遠距離交流」[40, 41] の成果であり、結果として、産業の構造転換（高付加価値化）が生じたものと捉えられる。

本論文では、新たな融合分野の創出を促す媒介者をネットワーク科学の手法を用いて分析する。具体的には、企業間のつながりのひとつである取引関係に着目し、取引をつなぎかえるリワイヤリングによってネットワーク構造を変化させる企業の解析を行う。取引関係がある企業間には、金銭や物品の交換だけではなく、取引を通じた知識や情報の交換も存在していることが想定される。ネットワークにおいて企業間のつながりが密なクラスターの内部に多数のつながりを保有する企業は、地域経済の活力に大きく影響を及ぼすハブとなっている可能性が高いが、ハブとは異なる意味でネットワーク内の重要な位置にいるのが、クラスター間をつなぐコネクターである。本論文では、コネクターの中でもネットワーク上で特に遠距離に位置するクラスター間を繋ぐブローカーとなる企業に着目する。

1.1.3 データサイエンスの導入によるネットワークの評価

(1) ネットワークを評価する既存指標と提案手法

ネットワークは点と、点を結ぶ線分から構成される。これらは研究者によって様々な用語が採用されており、ネットワーク内の点はノード、頂点、結節点、vertex と呼ばれ、線分はエッジ、枝、弦、弧、紐帯、辺、リンクとも呼ばれる [42]. ネットワーク内のつながりが密

な部分は、クラスター、コミュニティ、クリーク、モジュールとも呼ばれる。本論文ではネットワーク分析の手法を説明する際には点はノード、線分はエッジと呼び、エッジが密な部分を、企業間取引ネットワークではクラスター、社会ネットワークではコミュニティと呼ぶことにする。実社会の多くのネットワークはコミュニティ構造を有しており、コミュニティ内ではエッジが密であり、異なるコミュニティ間ではエッジが疎である。コミュニティとその特性を解析することにより、ネットワークのメゾレベルでの記述が可能になる。ここでメゾレベルとは、マイクロレベルであるノードと、マクロレベルであるネットワーク全体の、間のスケールを意味する [43]。

コミュニティ間をつなぐノードが有用な情報をもたらすことは、社会ネットワークの分析で着目され、定量的指標を用いた分析が試みられてきた。Granovetter は、頻繁に接するコミュニティ内の人々よりは、接する機会の少ない、コミュニティ外の人々によって有用な転職情報を得ることが多いことを発見し、これを「弱い紐帯 (weak ties)」の強さ論として体系化した [44]。これに対し、Burt はネットワーク内の構造的な位置付けを重視し、新規で異質性に富んだ情報はネットワークの「構造的な空隙 (structural holes)」を埋める相手から、効率的に取得できると考えた [45]。構造的な空隙を捉える尺度として Burt が考案したのが、構造的な空隙の欠如を表す「拘束度」である [46]。Burt の拘束度は 2 つのノード間の相互情報量と双方のノードの隣接ノードとの相互情報量を合わせた指標であり、コミュニティサイズが比較的小さなネットワークの構造的な空隙を捉えるのには有効である。しかし、定義上、注目するノードから 2 パス先までの情報しか捉えていないため、大規模な複雑ネットワークのコミュニティ間に存在する構造的な空隙は検出できないという限界がある。

一方、Guimera らが提案したノード指標は、ノードの所属コミュニティに対するエッジの分布の観点からネットワークの構造的な空隙を検出可能にしている。コミュニティ内部に多数のエッジを持つハブを識別する Z 指標 (within module degree) と、コミュニティ間にエッジを多く持つコネクターを識別する P 指標 (participation coefficient) である [47]。P 指標は Burt の拘束度では検出できない、複雑なネットワークにおいて構造的な空隙を埋めるノードを検出することができる点で優れている。しかし、P 指標が識別するのはノードの持つエッジがコミュニティの内部にあるか外部にあるかの違いでしかないので、コミュニティ間の関係性を捉えているとは言い難い。本論文では、P 指標にコミュニティ間の距離の概念を反映させることで、遠く離れた異質なコミュニティ間をつなぐブローカーの役割を果たすノードをより高く評価する指標 PW を提案する。

図 1.1 は、コミュニティ間距離が異なる 5 つのコミュニティを持つネットワークを用いて構造的な空隙を識別する 3 つの指標である P 値、PW 値、Burt の拘束度の逆数による評価を比較したものである。ネットワーク中のノードの緑色の濃さが各指標での値を表しており、数字は上位 10 ノードの順位である。ネットワーク中でエッジ密度の高い部分がコミュニティを表しており、エッジが少なく、コミュニティ間の距離が遠いところがネットワークの構造的な空隙にあたる。Burt の拘束度の逆数は構造的な空隙を捉えることができず、むしろコミュニティ内部のノードを高く評価していることがわかる。P 値はコミュニティ間の識別には成功しているが、距離が近いコミュニティ間をつなぐノードを高く評価している。これに対し、本論文の提案指標である PW 値は距離が遠いコミュニティ間をつなぐノード、すなわち構造

的空隙を埋めるノードの評価に成功していることがわかる。

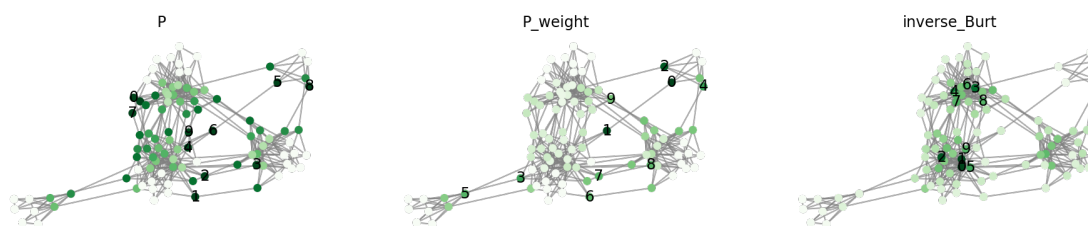


図 1.1: 構造的空隙を識別する 3 指標の比較. LFR ベンチマークネットワーク [48] を用いて筆者作成.

(2) サイバー空間上の情報の偏り

異質なコミュニティを結びつけるブローカーの重要性は、リアルな物理空間ではもちろん、サイバー空間上の人々のつながりのネットワークにおいても一層高まっている。サイバー空間では、コミュニティや発信側の意図に起因する情報の偏りが顕在化している。広告や商品推薦をはじめ、通常我々が目にするデジタル情報は過去の検索や閲覧履歴等に基づいて推定され自動的に決定されている。サイバー空間にはこの他にも、本人がお気に入りに入れたリンク先や発信者リストからの情報をはじめ、選別された特定種類の情報が続けて届く仕掛けが随所にある。ターゲットの嗜好や関心に合うよう選別された特定種類の情報が続けて届くフィルターバブル [49, 50]、閉じたコミュニティ内部での情報交換を繰り返すことで同種の意見が増幅・強化されるエコーチェンバー現象 [51, 52] など、限定された範囲でのやりとりで起因する情報の偏りが指摘されている。好みの情報にアクセスし、好みに合いそうな情報を推薦する仕組みの中では、自分やその所属するコミュニティとは異なる視点を取り入れることが従来以上に難しくなる課題が顕在化してきている [53]。情報技術の発達により、我々は世界中の膨大な情報を容易に入手できるようになったが、入手可能な情報量の爆発的な増加に対し、自分の考えの延長上にはない異質な情報に出会う機会は逆に減っているのである。

例えば、facebook は世界各国から多様な興味関心を持つ人々が利用しているが、そこでの情報取得は主に自分の友人もしくはその周辺の情報に限定されていることが多い。ユーザーの友人リストは卒業生、元クラスメート、故郷が同じ人、趣味のグループ、仕事仲間などさまざまなコミュニティに分けられ、各コミュニティ内の人間関係は密であり、コミュニティ間の人間関係は疎である。facebook 上で繋がっている複数ユーザーのエゴネットワークを統合すると、ユーザー間で共通の友人の数にばらつきがあることがわかる。図 1.2 は 10 人の facebook ユーザーのエゴネットワークを合体させて可視化したものである。ノードはユーザーの友人、エッジは友人間の facebook 上でのつながりであり、ユーザー自身は表示されていない。友人間のつながりが密なコミュニティはそれぞれ明るさの異なる灰色で表示しており、全ての友人はいずれか 1 つのコミュニティに所属している。橙色の星印で示している 10 人のユーザー間で共通の友人は、概してコミュニティの間に位置しているが、共通

の友人が多いコミュニティ間の距離は近く、共通の友人が少ないコミュニティ間の距離は遠いことがわかる。

友人を介して得られる情報の稀少性を考えるとき、距離が遠いコミュニティ間をつなぐ共通の友人の価値は高いと考えられる。直感的にわかりやすい例として、同窓会が挙げられる。同窓会で友人たちと話すのが楽しいのは、彼らと過去同じコミュニティに所属して過ごした時間があるからだけではなく、現在彼らが所属するコミュニティが、自分の所属コミュニティとは全く異なっているからである。本論文の第4章で提案するのは、ネットワークからこうした稀少な価値を持つ友人を見つけ出す指標である。



図 1.2: facebook 上の 10 人のユーザーの友人ネットワークと共通の友人. SNAP Dataset[54] を用いて筆者作成.

(3) データが駆動する地域創生の先駆的な事例

ネットワークの可視化と定量的解析は、地域経済関連施策の立案・実行・検証の上で重要な要素となっている。たとえば、経済産業省と内閣官房が提供している地域経済分析システム Regional Economy Society Analyzing System (RESAS) は、地域経済の産業構造や人口動態、企業活動等を可視化し、各自治体の政策担当者の助けとなる情報を提供している。梶川らはネットワークのメゾレベルでの特徴であるコミュニティ構造に基づくノード指標である Z 値と P 値を、企業間ネットワークに適用し、Z 値、P 値を軸とするマトリックスの分析により「コネクターハブ指標」を提案したが [55]、この指標は RESAS で公式に利用されているほか、2014 年の中小企業白書 [56] にも採録されている。2017 年 8 月に施行された「企業立地促進法の一部を改正する法律案」（通称：地域未来投資促進法案）では、地域経済を牽引する中堅企業（地域未来牽引企業）を重点ターゲットとして支援し、成長性の高い分野への集中投資が行われているが、コネクターハブ指標はその選定にも利用されている。

地域経済に好循環をもたらす企業を支援対象として選定する上で大規模なネットワークデータを活用することは、政府の政策支援効果を高める上でも重視されている。しかし、コ

ネクターハブ指標で用いられた P 値には、コミュニティの多様性を十分に反映できないという課題があった。本論文で提案する PW 指標では、取引の量ではなく質に着目することで、この課題を乗り越えている。この手法を用いれば、オープン化が進み、多様な知見や情報の融合が鍵を握るようになってきたイノベーション活動において、重要な役割を果たしている企業群の候補を抽出することが可能となる。また、専門分野や関心が異なり、ネットワーク上の距離が大きい学者グループや研究機関をつないでいる人物を見つけるなど、コミュニティに多様性を取り入れる仕組み作りへの様々な応用が考えられる。リアル空間、サイバー空間双方での情報交換も含めた、より柔軟な異分野間の連携を誘発するネットワークの動態や構造、コミュニティ形成のプロセスを明らかにすることは、今後一層重要になることが見込まれる。そこでは地理的近接性といった距離の概念や、産業分類・地域といった従来の集積単位を多様化し、ネットワークの観点から再構築していくことが求められるだろう。

1.2 着想の原点と本論文のアプローチ

前節では本論文で取り組む基本的な 2 つのテーマ、すなわち、変動する市場環境に対応する企業の新陳代謝と、ネットワークの構造的空隙を埋める稀少なつながりに関する社会的背景や課題、政策への生かされ方を具体例とともに述べた。本節では、本論文の着想の原点となった考え方とアプローチについて概説する。

(1) 生物と経済の接合過程からの着想

Schumpeter は「企業組織上の発展は、不断に古きものを破壊し、新しきものを創造して、たえず内部から経済構造を革命化する産業上の突然変異によってなされていく」とし、企業の非連続な動的発展を突然変異という生物特有のダイナミクスになぞらえて捉えて、創造的破壊と表現した [23]。産業集積論やイノベーション研究では、生態系を意味する用語であるエコシステムという言葉がよく使われるように、本論文のアイデアも、生物学的ダイナミクスと経済活動のアナロジーに着想の原点がある。生態系は生物個体間の関係性の絶妙なバランスによって平衡状態に近づこうとする動的システムである。企業間ネットワークにおける個別企業のミクロな動態と他企業との関係性から、地域経済の望ましいあり方を考えるとき、生命体や生態系の持つ柔軟な秩序形成機構は参考になることが多い。

Iansiti らはビジネスネットワークの概念化に最も有効なのは、生物界のエコシステムから多くを学ぶことであるとしている [57]。本論文第 2 章では、生物界のエコシステムの具体例として極相林を取り上げ、極相林を平衡状態に近付ける動的機構であるギャップダイナミクスの概念を紹介する。「ギャップ (gap)」とは落雷や倒木などにより森林に生じる穴のことである。近年の森林生態学では極相林はギャップの形成と閉鎖を通じて、種多様度が保持されるという考えが重視されている [1]。Iansiti らは、エコシステムの健全性をもたらすのは多数の企業と相互作用を持つキーストーンであるとし、よいキーストーンとなる企業の条件として、新しい「ニッチ (niche)」を創り出す能力を挙げた。ニッチはエコシステムの周辺に位置し、容易に入れ替わるゆるやかな連結で異なる複数のエコシステムをつなげる

場となる [57]. 近年の経済学における類似概念としては, Christensen が提案した「無消費 (nonconsumption)」があげられる。「無消費」とは消費するモノやサービスがない状態のことである。Christensen は無消費経済に集中することは企業の新たな成長エンジンに点火する機会となりうるとし, そうした新たな市場創造が持続的経済成長を可能にするとした [58]. ギャップ, ニッチ, 無消費は, いずれもエコシステムの「穴」に相当する概念であり, 創造的破壊は既存の経済均衡や秩序を破壊して市場に穴を作る行為と言えるが, 極相林のような植物遷移の最終段階においても, 穴の開閉による動的秩序形成機構が見られることは示唆深い。経済発展の成熟段階にある日本のビジネスエコシステムを考える上でも, 企業間ネットワークの構造的空隙の分析は, 有益であることが推察される。

(2) 柔軟性を維持するネットワークの構造と推移の解析

本論文では, ネットワークを構成する要素間の関係の動的推移と構造特性を捉える 2 つのノード指標を提案する。1 つは隣接ノード群の平均入替率を表す指標: 新陳代謝度であり, もう一つはエッジの接続先コミュニティの異質性を表す指標: 重み付き P 値 (PW) である。新陳代謝度はノードが保有するエッジの動的推移解析に用いられ, PW 指標はノードのネットワークにおける位置に関する静的構造解析に用いられる。隣接ノードを頻繁に入れ替えるノードは新陳代謝度が高くなり, 遠く離れたコミュニティないしクラスターの間隙にあたる, ネットワークの構造的空隙を埋めるノードは PW 値が高くなる。指標の妥当性を検証するため, ノード数, エッジ数, コミュニティサイズ, リワイヤリング率等の条件を変えて生成した複数のベンチマークネットワークを用いた。また, 実データとして facebook データ, 企業間取引データを用いた分析を行い, 提案指標によるノード評価について考察した。

2 つの提案指標はいずれも個別のノードのネットワーク内における評価を行うものであり, 複数のノードがクラスターのような集団としてネットワークでどのような位置にあり, 推移するのは評価できない。そこで次に 10 年分の企業間取引データを用いて提案指標に基づくノード評価とクラスターの推移の関係を分析した。その結果, ネットワークの構造的空隙を埋める PW 値上位の企業の傾向として, 取引先企業が少ない, 移り変わりが激しい, クラスター間のモビリティが高いという 3 つの特徴があることがわかった。また, PW 値の高いノードはクラスター継承性の予測に寄与することがわかった。

本論文のリサーチクエストは 3 つある。一つ目は, ネットワークの動的推移に関する問いである。相互に結びついたビジネスエコシステムである企業間ネットワークにおいては, 他社との適切な関係構築が企業の生存戦略上の大きな柱となる。しかし, 今日のように変動の大きい市場環境下では, 企業間関係を安定的に維持するだけでなく, 柔軟に変更できる能力が, 企業の存続にとって重要になっている可能性が高い。そこで, 他の企業との取引関係をどのように推移させる企業が, 地域市場で長く存続できるかに着目して分析を行う。

二つ目は, ネットワークの構造特性に関する問いである。先行研究からは, 信頼性の高い企業間関係の継続とともに, 異質な情報や知見の流入が, イノベーションを誘発する上で重要であることが示されている。そこで, 企業間ネットワークにおいて, 異質性の高いクラスター間をつなぎ, 異質な情報や知見を自社が所属するクラスター内に取り込む役割を果たし

ている企業群（本論文において「ブローカー」と呼ぶ）を抽出し、それらの企業がどのような特徴を有しているのかを明らかにする。なお、本研究における「ブローカー」とは、ネットワークの構造情報に基づいて算出したPW値上位5%以内に入るノードのことを指し、金融・証券用語として用いられる売買の仲介人のことは意味しない。

三つ目は、企業間ネットワークにおけるクラスターの推移に関する問いである。クラスターの進化イベントを識別する既存手法では、ネットワーク内でどのようなポジショニング戦略を取るノードが、クラスター進化に影響を及ぼすのかが、定量的には明らかになっていない。そこで、提案指標を含むノードのポジション価値を表すネットワーク指標を用いて、クラスター進化の予測を行い、各指標の予測への寄与度を評価する。また、進化イベントをはじめとするクラスターの推移パターン（本論文において「クラスターパス」と呼ぶ）を抽出してその特徴を分析し、ブローカーがクラスター内で果たす役割について検証する。

1.3 貢献

本論文の学術的独自性は、理論的には想定されていた企業レベルの新陳代謝やネットワークの構造的空隙をつなぐブローカーを検証可能な形で新たに定義したことにある。これにより、大規模データを用いて企業の経済活動のダイナミクスを定量的に解析することに成功した。本論文の貢献は以下の通りである。

- 企業の新陳代謝度をノードが保有するエッジのつなぎ替え率として定義し、大規模データに適用して日本企業の新陳代謝を初めて定量的に測定した。また、企業の存続年数と新陳代謝の関係を明らかにした。
- 不均質なコミュニティ構造を持つネットワークにおける、エッジの稀少性を評価するノード指標PWを開発し、これを用いて検出されたノードのネットワーク内での重要性を定量的に検証した。
- 開発した上記2つの指標を実際の企業間取引ネットワークに適用し、クラスター進化の観点から新たな評価を行い、PW指標がクラスター継承度の予測に最も大きく寄与することを明らかにした。
- クラスターパスの特徴を定量的に分析しケーススタディを行うとともに、それぞれの特徴に応じた政策的支援のあり方を提示した。

1.4 本論文の構成

第2章では経営戦略における企業ネットワーク研究、生物のエコシステムからの着想に基づく持続的社会モデルへの試案、ネットワーク科学の分析手法に関する関連研究と本論文の位置付けについて述べる。第3章ではノードのリンク組み換え率を用いたノード評価指標（新陳代謝度）によるネットワークの推移解析について提案する。第4章ではノードが媒介するコミュニティ異質性を用いたノード評価指標（PW値）によるネットワークの構造解析

について提案する。第5章では提案手法について実データへの適用および、クラスター進化の観点でのケーススタディを行い、その結果について述べる。第6章では、実験結果に関連する類似概念のまとめと考察を行い、提案手法の適用可能性と課題について議論する。第7章では本論文の結論と展望を述べる。

提案指標の位置付けを含む本論文の枠組みを表 1.1 に示す。

表 1.1: 本論文の枠組み

	動的・推移解析 (3章)	静的・構造解析 (4章)
企業間の取引関係 (ミクロ)	指標：新陳代謝度 / 適応性 対象企業が取引先の増減や入替によりどれだけ取引先構成を変化させたか	指標：PW 値 / 稀少性 対象企業の取引先がネットワーク全体の取引先構成の多様化にどれほど寄与しているか
クラスター間の関係 (マクロ)	指標：クラスター継承度 / 群間変動・進化 (5章) 連続する2年に存在するクラスターから構成したクラスターペアにおける企業群の類似度	

第2章 関連研究および手法の着想

本章では、地域クラスターと企業間ネットワークに関する先行研究を振り返り、本論文における着想の契機となった生物学的事象とのアナロジーとして極相林のエコシステムと新陳代謝に関する概念を説明する。また、関連するネットワーク科学の分析手法として動的推移解析と静的構造解析、コミュニティ進化イベントの検出と予測について説明する。最後に本論文の位置付けについて述べる。

2.1 地域クラスターと企業ネットワーク研究

本節では地域に集積した企業間の相互作用や集積の利点を踏まえたイノベーションモデル (TIMs: Territorial innovation models) に関する研究を概説し、近年の傾向として地域を超えて進化するネットワークのダイナミクスに対する関心が高まっていることを示す。また、企業が所属するコミュニティを離れた外部知識を吸収するために必要なつながりの多様性に関する研究について論ずる。

2.1.1 地域クラスター戦略：産業集積の利点と課題

(1) 地理的近接性と集積効果の限界

TIMs は地域クラスター (regional cluster)、産業地帯 (industrial districts)、イノベティブ・ミリュー (innovative milieu) を含む包括的概念であり、地域規模の経済活動の集積と、その中に位置する企業にもたらされると思われる利益に着目する伝統的イノベーションモデルである [59]。TIMs は企業間ネットワークに持続的な競争上の優位性をもたらす産業構造について論じており、地理的に近接した地域に位置する企業は、互いに協力関係・競争関係を得やすいため、企業間の相乗効果を促進すると考えられてきた [30, 31, 32, 33]。

Porter は地域クラスターの利点として、顧客獲得と保持のための競争を生じると同時に関連企業間の協力を促進すること、アライアンスやパートナーシップといった非公式な関係構築により、ネットワークの効率や柔軟性を高めるなどの点を挙げている [32]。こうした利点は地域クラスター内の知識交流やイノベーションを促進し、生産性や成長率の向上に結びつくとしてきた [60]。たとえば、Saxenian はシリコンバレーに集積したハイテク産業を調査し、非公式な情報交換も含めた企業間の柔軟な関係性が、製造サイクルの短期化や急激な技術進歩といった環境変動に呼応する上で重要な役割を果たしたことを明らかにしている [61]。

一方で、地域クラスターに特徴付けられるような地理的近接性は、必ずしも企業間の相互作用を促進するとは限らないとの報告もある [62, 63, 64, 60]。Amin はグローバリゼーションが進行する中、企業間の交流手段は多数存在するため、地理的近接性に限定されない、ネットワークを構成するアクターや、アクター間の関係構築のプロセスへ重点を移行することが必要だと述べている [62]。Sheppard は「距離」の概念には経済的距離、社会的距離など様々なものがあるが、地理的経済分析では2点間を結ぶ直線の物理的長さであるユークリッド距離ばかりが重視されてきたことを指摘し、多様な距離の概念や、結合を生じさせるプロセスを考慮する必要があると論じている [63]。Boshma は近接には地理的近接以外にも、認知的近接、組織的近接、社会的近接、制度的近接という異なる次元があるとした上で、近接にはロックイン効果などイノベーションに負の影響を与える可能性もあると指摘し、近すぎず遠すぎない効果的な距離の調整や柔軟性の確保について述べている [64]。これらの指摘はいずれも、近接性や距離の概念の拡張といった企業間交流を促進する条件の再検討と、ネットワーク形成プロセスへ着目と捉えられる。情報技術の進歩とグローバリゼーションが進んだ現代市場において企業間交流の手段は多様化し、相互作用を促進する要因としての地理的近接性の重要度は、相対的に低くなってきていることが伺える。

こうした距離概念の拡張や動的推移分析の課題に加え、集積効果の利点に偏った分析に対する懸念もあげられている。Martin は、近年の都市経済モデルでは、都市がほとんど「自由に浮かんでいる島 (free-floating islands)」のように扱われ、都市のロケーションは集積効果ほどは重視されていないとし、地域クラスターの生成・進化過程や地域クラスター間の競争・相互作用を、より広い経済システムに結びつけた研究が不足していると指摘している [60]。Porter は地域クラスター戦略が成功するための重要な留意点として、ローカルな競争環境下で担保されるクラスターの更新 (cluster upgrading) をあげているが [31, 32]、それはあくまでクラスターを構成する企業の技術進歩や市場開拓といった努力によって、所属クラスターが更新されることであり、複数クラスター間の相互作用を積極的に捉えたものではない。しかし、特定地域や産業における集積効果が大きかったのは、今日のように異分野融合が活発化する前のステージであったと言える。Porter が着目したのは、1990~2000年の米国における、特定技術を中心とする分野特定型クラスターの集積効果であったが [30]、近年の先進国市場においては、既存技術の漸進的成長をもたらすクラスター内部の集積効果に加え、異なるクラスターをつなぐワイヤリングによりもたらされる異分野融合が一層重要になってきている [40, 41]。

(2) 変動する市場環境で求められる外部知識の吸収力

地域クラスター戦略をはじめとする TIMs が前提としていた、重要な情報は常に特定地域の内部で交換されているという想定は近年疑問視されている [60]。企業間コミュニケーションの充実や地方政府による支援といった産業集積の優位性が行き過ぎると、イノベーションには障害となる硬直的な専門化、すなわち社会経済的なロックイン状態に陥るとの指摘もある [65]。新しいイノベーションの機会を探るためには、多様な技術や知識源に触れることが不可欠であるが [66]、地理的近接性を必要としない知識の溢出 (spillover) に関する研究も多

くある [64, 67, 68, 69, 70]. Cohen と Levinthal は新しい外的な情報 (external information) の価値を認識して吸収し、ビジネスに結びつける企業能力がイノベーションに欠かせないとし、これを吸収力 (absorptive capacity) と呼んだ [5]. 外部性への着目は、企業の周辺部からの弱いシグナル (weak signals from the periphery) [71, 72] や、非冗長なアライアンス・パートナー (non-redundant alliance partners) [73] といった概念にも見られる。

外部から技術や知識の供給を受けつつ、自社の得意とする分野に特化した技術更新に集中することは、最先端の技術水準を保つことを可能にするが [61], 外部知識には自社にはない他社からの知識 (external knowledge) に加えて、所属クラスターの外からの知識 (extra-cluster knowledge) も含まれる [74]. O'Reilly は、製品やサービスのモジュール化が進み、通信コストが低下するにつれ、イノベーションの場はますますクラスターに移っていくであろうと述べ、今後の探索と開発に関する研究は、企業内や企業間に焦点を当てたものから、より大きなクラスターを意識したものへと移行していかなければならないだろうと述べている [75]. 特に技術進歩が著しい産業においては、イノベーションを促進する異質な知識は、クラスター内部ではなく、クラスターから離れた位置にある可能性がある。例えば、ハイテク企業の近年の分析は、企業の吸収力と関連しているのは、地理的近接性よりは取引先企業との関係性の近さであること、クラスターから遠く離れた場所に位置する革新的で活発な顧客との関係による多様性により、イノベーションの機会へのアクセスが改善することを示している [76].

遠距離からの知識 (distant knowledge) を取り入れることで生じる知識の不均一性は、企業間協力の多様化や、変動する市場環境に対応する組織更新への有効性の観点から評価されてきた [77, 78, 79, 80, 81]. ローカルミリューを離れた外部環境との接点、すなわち、技術や市場、経営に関する自社が持たない、異質で多様な知識への接続と相互交換は、大企業のみならず中小企業のイノベーション過程にも見られる特徴であることが判明している [82, 83, 84]. 企業間ネットワークのリワイヤリングによってもたらされる、遠距離の異質なビジネス相手は、企業が新興技術の脅威を認識したり、イノベーション・ソリューションとなる技術を開発するのに役立つ新しい情報をもたらす可能性がある [40, 41].

2.1.2 地域ネットワーク研究：つながりの価値とブローカーの役割

(1) ポジショニング戦略論と内部資源論

企業を取り巻く外部環境を分析し、産業構造に応じた自社独自のポジションを取ることが持続的な競争優位をもたらすとした Porter の競争戦略論 [85] に対し、企業内部の経営資源に競争優位の源泉を求めたのが、内部資源論 (RBV=Resource Based View) である [86, 87]. RBV 理論は産業を構成する企業群が保有する資源の異質性 (resource heterogeneity) に着目し、それらの資源は企業間で容易には移転されないこと (imperfect mobility) を前提条件として、そうした異質性の維持が企業の持続的競争優位につながるとしている [88, 89]. RBV は他社が真似できない企業独自の能力であるコアコンピタンス (core competence) 戦略を支持している。Prahalad と Hamel はコアコンピタンスを他社が真似できない理由として、

それが単なる個別技術の集合 (collection of discrete business) ではなく、技術の複雑な調和によるものであり、組織内調整と学習を含んだ包括的パターンを有することを挙げている [90].

ポジショニング理論は企業の外部環境, RBV 理論は企業の内部環境に着目し, 互いにもう一方は前提としている点で相互補完的である. しかし, これらは競争環境を比較的安定したものと捉える静態的理論であり, 技術進歩や製品ライフサイクルが速く, 顧客ニーズが多様化した現代の市場への適用には限界がある. 変動が激しい現代市場に即した戦略立案には, 市場における企業のポジションと内部資源双方の変動を踏まえた分析が必要となる [3].

(2) ネットワークでつながった企業の価値

RBV 理論を拡張し, 企業間のつながりを知識媒体としての資源と捉えたのがネットワーク資源論 (network resources) である [91, 92]. 個々の企業内での内部資源の蓄積と差別化に焦点を当てる伝統的な資源戦略論 [88, 89] とは対照的に, ネットワーク資源論では, ビジネス上のつながりから得られる外部情報も含めて資源とみなす. 企業間連携における提携先企業の資源プロファイルの分析結果は, 提携先企業の特徴に応じて利益の発生率が異なることを明らかにしている [93, 94]. Lavie はつながり合った企業が互いの資源からどのように価値を引き出すかを分析し, 競争優位は企業内部の資源だけではなく, 企業間の資源の相補性によってもたらされるとした [92]. Bellamy らは取引ネットワークのアクセス性と相互接続性に加え, 企業の吸収力と取引先企業の革新性が, 企業のイノベーションにとって重要であること, 相互に結合したネットワークの中にある企業はイノベーションを起こしやすいことを示している [95]. Powell は専門知識が広く分散している場合には, イノベーションの源泉は組織間学習のネットワークの中に見られると論じている [96].

ネットワーク特性は企業パフォーマンスに重要な影響をもたらすが, 同質な相手との結びつきを強くすればイノベーションが盛んになるというわけではない [97]. ネットワークでつながっていることに価値を見出すネットワーク資源論では, 他社との緊密な関係や長期的な関係を持つことが信頼性の醸成などの利点につながるとする研究が多く見られるが [98, 99, 100], 逆にネットワークへの埋め込み (embeddedness) が行き過ぎると硬直化してイノベーションを生みだしにくくなるとの指摘もある [99]. Powell はどういった場合にネットワークが硬直化しやすいのかに着目し, 企業発展の初期段階では運営統合や, 規模の経済, リスク回避に有利だった垂直統合の強みが, 企業発展の後期段階では硬直化を生み, 弱みになる場合があるとした. また, 垂直統合型の大企業が減少し, 中間型の企業が増加傾向にあることを指摘し, その要因として国際市場での変化に迅速に対応できなかったこと, 既存の製造過程における取引関係の改革や新製品の実装への抵抗があったことをあげている [101]. 過度な強連結・凝集的なネットワークは硬直化しやすく, 革新しづらくなってしまいう機能不全現象を起こしやすい.

サプライチェーンの柔軟性 (flexibility) や鋭敏性 (agility) が, 企業の成長や市場シェア, 技術の探索と深化といった企業パフォーマンスによい影響を与えることは, 多くの研究で指摘されている [102, 103, 104, 105, 106]. Yusuf らは, 鋭敏なサプライチェーンは, 激動するビ

ビジネス環境において、鋭敏性の低い競合他社よりも優れていることを実証した [4, 107, 108]. Hohenstein はサプライチェーンの鋭敏性が高まるほど、震災復興や企業パフォーマンスの向上は早まると述べている [10]. サプライチェーンに代替可能な取引相手の確保を可能にするネットワークの冗長性と柔軟性を組み込むことにより、企業は迅速な回復能力 (resilience) を獲得できる [7].

知識の専門化と多様化が進むにつれ、外部知識をもたらす緩やかなネットワークのつながり (loosely coupled network) はより重要になる. 特に多くの技術を扱うマルチテクノロジー企業では、依存する技術の開発速度のムラや製品レベルの予測不可能な相互依存関係によって引き起こされる不均衡に対処するために、自社での製造に要する知識の範囲に留まらない、多くの知識が必要となる [109]. 競争圧力は市場における階層構造の段階を減らし、新たな協力関係の探索と組織の境界の再定義を誘発する. 特に企業や大学、研究室、供給者や顧客といったアウトサイダーとの結びつきはイノベーションに重要である [101]. Chesbrough は、オープン・イノベーション・パラダイムにおいては、外的知識は取得済みの概念である内的知識と同等の役割を果たすと述べている. 企業はもはや単独でイノベーションを起こすのではなく、社内外のリソースから知識を得ることでイノベーションを起こすようになった. このため企業には、企業内部で知識を管理する能力に加え、情報ストックになりうる資産を活用して知識を獲得し、伝搬する能力が必要とされている [110, 111].

(3) 強いつながりと弱いつながりのバランス

強いつながりを持つことは、必ずしも多くのイノベーションが勃興することを意味しない [112]. 実際、企業間ネットワークにおける事業戦略を構築する際には、弱いつながりと強いつながりのバランスを気にする必要がある [113, 80, 98, 114]. Teece らは企業に競争上の優位をもたらす能力や資源の組み合わせがどのように開発・活用・保護されるのかに着目し、変動する環境に対応するために、社内外の能力を統合・構築・再構成する企業特有の能力としてダイナミック・ケイパビリティ (dynamic capabilities) の概念を提唱した [3]. ダイナミック・ケイパビリティは、ポジショニング戦略論と内部資源論を発展的に統合した概念であり、企業が自社の市場における位置 (market position) と、能力・資源の発展段階 (path dependencies) に応じて、イノベーティブな競争上の優位を更新していく動的プロセスに着目している. この考え方は変動する企業間ネットワークの中で、企業がどの程度、外部知識を取り入れて学習するのが経営戦略として最適かという次の課題につながる.

Powell は企業の生存率を高めるネットワークの埋め込みの正の効果は、ある閾値を超えると反転するとした [101]. では、ネットワークの埋め込み (強いつながり) とそれを緩和するネットワークのつなぎ替え (弱いつながり) は、どのようなバランスで行えばいいのであろうか. March は組織学習において実績の深化・活用 (exploitation) は短期的利益をもたらすが、探索・開拓 (exploration) に取り組まない企業は変化に直面した時に、破綻する可能性が高いことを示した. March が分析したのは、エッジのつなぎ替え (新規採用や転職)、組織の学習率、環境変動率の3つのパラメーターと組織のナレッジ水準との関係である. すなわち、新規採用や転職により構成員が入れ替わる時、

組織内の人がほぼ学習をしない場合には組織が保有するコアナレッジは減少するが、学習率が高ければ0.1-0.4程度の入れ替えがある時に組織ナレッジは高い水準になることがわかった。また、環境の変動率が大きくなると、構成員の入れ替えを全くしない組織のナレッジは劇的に低下するのに対し、入れ替え率が0.1の時には環境変動の影響をほぼ受けずに高い水準を維持することがシミュレーションにより示された [80]。

組織学習における探索と深化には最適なバランスがあることを示した March の研究は、企業活動において漸進的なイノベーションと不連続なイノベーションを同時に実現する両利きの経営のアイデアにも関連している [115]。提案者の Tushman と O'Reilly は後に、探索と深化のバランスを取ることの難しさは、企業が短期的成功の確実性が高い深化の方を選択しやすいことにあると指摘している。探索は非効率であり、筋の悪いアイデアの増加につながりやすいが、探索の努力をしなければ変化するビジネス環境への適応に失敗する可能性が高まる [75]。

(4) イノベーションを仲介するブローカー

Granovetter は人間関係のネットワークを通じた転職について調べ、頻繁に接するコミュニティ内の人々よりは、比較的關係が薄く接する機会の少ない、コミュニティ外の人々によって職を得ることが多いことを発見し、これを「弱い紐帯 (weak ties)」の強さ論として体系化した [44]。Granovetter が弱い紐帯、すなわち滅多に会わない相手とのつながりを異質な情報の取得機会として重視したのに対し [44]、Burt は異質性をもたらす構造に着目し、隔たった複数の相手とのつながりによって、新規で異質性に富んだ情報を効率的に取得できると考えた。Granovetter が見ていた「弱さ」の尺度は接触頻度であり、ネットワーク分析ではエッジの重さに相当するが、Burt は紐帯の弱さはエッジの相関事象にすぎないとし、重要なのはそれがコミュニティ間をつなぐブリッジであるかどうか、すなわちネットワークの隙間を埋めるかどうかであるとして、「構造的空隙 (structural holes)」論を展開した [45]。

Lester と Piore が「新しいマネジメント」と呼んだ企業の行動原理は、その中心的な目的を、企業経営の柔軟性を高め、経営環境の予知できない急速な変化への対応スピードを加速させることとしている [116]。柔軟性向上の主要な鍵となるのは、効率的な組織統合のあり方である。そこではいかに多様な人々や事業部門、取引先や顧客、他社との境界を超えて横断的な連携を生み出すかが重視される。彼らは最もよく知られている新しいマネジメントの規範として「顧客の声に耳を傾けよ」「得意とするコア・コンピタンスに集中せよ」という二つの命題をあげ、これらは既知の問題解決を志向する一元的な「分析的思考」であるとした。一方で、問題そのものを見つけ出すためには、異なる人々や組織の曖昧さを含んだ相互作用（対話）に基づいて、クリエイティブな関連づけを行う「解釈的思考」が必要であるとした。これを企業間ネットワークに当てはめて考えると、解釈的思考は企業間の相互関係が存在しない構造上の穴 (structural holes)[45]を埋める働きと見なせる。Lester と Piore は構造的空隙にブリッジをかけることにより、異なる主体間に共通の言葉と語彙が発達することが期待できると考えた [116]。構造的空隙には情報ギャップがあるため、これを活用することにより、利益を得る機会がある。構造的空隙を埋める仲介者は、単に知識を集めるだけでは

なく、企業内外の変動や改善を統合し、組織や人々の乖離や境界をつなぐ役割を担うことが期待される [109]. Huggins は企業にとって最良のネットワーク支援は、製品や市場開発を共同で行うハードなビジネス関係と、共通の問題を解決する非公式な協力関係であるソフトなつながりによる社会的利益を、ミックスして重ね合わせることができるブローカーによってなされるとした [117].

イノベーションがよりオープンになるにしたがい、企業間ネットワーク内の仲介者であるブローカーは、イノベーションの形成過程に一層重要な役割を果たす [113, 118]. オープンイノベーションの提唱者である Chesbrough は、イノベーションを促進するために異なる企業をつなぎ、アイデアを媒介する役割を担うことで双方に有益な知識移転を実現する、イノベーション仲介企業 (innovation intermediary) の概念を提案している [119, 120, 79]. イノベーション仲介企業はイノベーターが社外のアイデアを活用するのを援助したり、発明者が自身のアイデアを他社に利用してもらうための市場を見つけて、相互利益を得ることを援助し、イノベーションのプロセスをオープンにすることに貢献している [119]. イノベーション仲介企業には、小売業者、メディア企業、広告代理店、市場調査機関などが含まれ、不確実性の高い新興市場で技術を仲介し、一方の企業の代理人として機能したり、2社を引き合わせて取引を実行することでマーケット形成を促進する役割を担っている [29]. しかし、イノベーションのプロセスにおける仲介者の特徴を調査した研究は、イノベーションの新規性に注目する研究の蓄積に比べて少ない [120].

2.1.3 地域クラスターと企業間相互作用のダイナミクス

ここまで、地域クラスターをはじめとする、集積の地域性に着目するイノベーションモデルである TIMs の概要と課題を踏まえつつ、地域クラスターの構成要素である企業と、企業間をつなぐネットワークに関する研究の流れを述べてきた。本節では、企業間相互作用のミクロな変動が、地域クラスターのようなマクロなネットワーク構造のダイナミクスにどのような影響を及ぼすかに着目し、旧来のクラスター研究の限界に言及する。また、TIMs を拡張したモデルとして、地域外の知識媒体との動的相互作用に着目し、クラスターのダイナミクスを捉える TKDs (Territorial Knowledge Dynamics) について論ずる。さらに、企業間競争ではなく、エコシステムとしての企業ネットワーク内の企業の役割分担に焦点を当てたキーストーン戦略とその課題について述べる。

(1) ミクロな駆動力とマクロな構造変化

地域産業のパフォーマンスは、しばしば地域に集積した企業の相互作用の変動、すなわち地域クラスターのダイナミクスの観点で解析されてきた [30, 121, 33, 122]. 地域クラスターの利点の一つは、知識とイノベーションの媒介者としての役割を果たす、地域の企業間ネットワークの形成を促進することである [123, 124]. しかし、Porter は地域クラスターの利点として企業間連携の促進を挙げる一方で、クラスターの萎縮や衰退の原因は、クラスターを構成する企業間関係の硬直化と、外部環境との接続不足によるとした。規則強化や相互依

存、カルテルといった競争障壁によるクラスター内部の過度な硬直は、イノベーションの速度低下をもたらし、外部環境からの遮断は先進技術への接続を困難にする。このため、クラスターの利点を享受するには、クラスターの更新を促す競争的なビジネス環境が不可欠であると指摘した [32].

しかし、企業間ネットワークの中のクラスターは、その特徴や進化の道筋、イノベーションへのアプローチなどが異なり、アウトプットに影響を与えている [67, 125]. ネットワークのマクロな構造的特性の進化は、ミクロレベルで進行する力によって駆動されることが示唆されているが [124, 126], 個別の企業による取引先の変更をはじめとするミクロレベルの変動が、マクロな地域クラスターの変動傾向と常に一致するとは限らない。近年の地域クラスターのダイナミクス分析では、企業間に新しいつながりが多数出来、ミクロレベルで大きな変動があっても、巨視的なネットワーク構造は安定していることが判明している [127].

Martin はクラスターの状況と、クラスターを構成する企業間の競争や相互作用を、より広いグローバル経済システムの発展と結びつけている研究がほとんどないことが、クラスター政策を考える上で最大の問題だとしている [60]. Wal と Boschma はクラスター研究が陥りやすい欠点として、(1) クラスターを構成する企業の能力、特に外部知識を効果的に吸収する動的能力が異質であることを無視していること、(2) 地理的近接性の重要性を過度に強調し、ネットワークの役割を過小評価していること、(3) クラスターの起源と進化について考慮していないことをあげている。彼らのアプローチは、静的な分析に偏りがちであったクラスター研究の限界を示しつつ、それを乗り越える新たなフレームワークとして、ミクロレベルでの企業、マクロレベルでの産業、そして両者の相互作用を表すネットワークの3つが共進化する形で、クラスタリングのパターンが現れることに着目するものであった [128].

(2) 知識係留と多地点ネットワークの進化モデル

イノベーションは、企業間の知識の流れを必要とする複雑な相互作用を通じて形成される [110, 91, 88]. 知識や情報、人、資金といった生産要素の移動性 (mobility) の増大と知識係留 (knowledge anchoring) のダイナミクスに着目した Crevoisier らは、集積の地域性に着目した TIMs を拡張し、地域の境界を超えて進化する知識ネットワークである TKDs (Territorial Knowledge Dynamics) の概念を提案した。TIMs における学習プロセスが、主にローカル／グローバルといった地域の規模に応じて捉える累積的な知識のダイナミクスであったのに対し、TKDs では学習プロセスを、知識の移動性と係留性を反映した多地点 (multi-location) ネットワークで捉える知識の組み合わせ (combinatorial knowledge) のダイナミクスとして発展させている。TKDs では、知識移動の担い手と移動先の新しい環境とは相互に変化する依存関係を持ち、学習効率は知識係留の強度に依存する。すなわち、知識係留の強度が強ければ、知識移動の担い手や移動先の知識は豊かになる [129].

James らは、地域を超えた相互作用の中で知識が生まれ、組み込まれ、循環する知識係留のアイデアに着目し、TKDs の再解釈とケーススタディを行っている [130]. 彼らは暗黙知のような重要な知識の交換は常に地理的近接性が優位に働く face-to-face の交流の中で生まれるとする TIMs の想定は、近年疑問視されているとし、イノベティブな企業が知識交流

を求めて頻繁に越境することがこれまでの実証研究で裏付けられていることを示している。James らによる調査はスウェーデンの地方の自動車部門のネットワークにおけるキーパーソン 15 人を対象にした比較的小規模なインタビューに基づくものであったが、抽象的概念に留まっていた TKDs を実際のデータに基づいて検証している点で意義深い。また、ネットワークの規模、構成、相互作用、変化の速度といった観点で地域ネットワークを捉える彼らのフレームワークは、TKDs の実証研究を運用可能にしている [130]。今後、より大規模なパネルデータと、検証可能な指標に基づくネットワーク分析手法を導入することにより、TKDs 研究は精緻化されていくことが見込まれる。

(3) ビジネスエコシステムにおけるキーストーン

競争の少ない環境を選択することの利点に着目したポジショニング戦略論 [85] や、他社が真似できない企業独自の能力（コアコンピタンス）に焦点を当てた RBV 理論、そして地域クラスター戦略といった企業間の競争優位をベースに論じる戦略モデルに対し、Iansiti らは、決定的な競争は企業間ではなく、企業ネットワーク間で起きているとし、運命を共有する複数の企業からなるエコシステムとしての生存戦略を提唱した [131, 57]。

ここで興味深いのは、ビジネスネットワークの概念化に最も有効なのは、生物界のエコシステムから多くを学ぶことだとしている点である。Iansiti らは外部からの衝撃に対する堅牢性、専門化と目新しさを創出する能力、異種混合の結合といった生物界の特徴に着目して、ビジネスエコシステムを捉えた。そして、エコシステムの健全性をもたらすのは多数の企業と相互作用を持つネットワークのハブを専有するキーストーンであるとし、その特徴の一つとして新しいニッチを創出する能力を上げている。有能なキーストーンはビジネスエコシステム全体の健全性を高めるために行動することになるが、そのためには外部からの刺激を排除するのではなく、むしろそれをエネルギーとして新しいニッチを創出し、エコシステムの多様性を高める方向に転換しなければいけないとしている。また、ニッチを活動の場とするニッチプレイヤーは、エコシステムの周辺に位置し、異なる複数のエコシステムをつなげることができる点、容易に入れ替え可能な緩やかな連結により、エコシステムの変動に対応できる点が特徴とされている。

エコシステムを広義のクラスターと捉えるなら、これは企業間の競争によってクラスターを更新し、企業間関係の硬直化や外部環境からの遮断を防ぐという、ある種自社中心主義の企業行動がクラスター全体の維持発展につながるとする Porter の観点とは異なり、クラスター全体の維持発展をもたらす企業行動が各企業の利益にもなるとする考え方である。企業の異質な能力に基づいて個別の役割を認めている点、地理的近接性よりは企業ネットワーク内の企業間相互作用を重視している点で、Wal と Boschma の (1)(2) の指摘を乗り越えるアプローチともいえる。しかし、エコシステムの起源や進化を考慮していない点で Wal と Boschma の (3) の指摘は未解決のままである。また、具体的にネットワークがどのような条件にある時にニッチが創出されるのかは明確には示されていない。Iansiti らはニッチ創出の適切で直接的な測定方法は、新しい技術が新事業や新製品の多様な形態にどの程度まで転換されているかを見ることとしているが、これは結果論であり、ニッチを検出するための手が

かりにはならない。本論文第4章では、異なるエコシステム（クラスター）をつなぐ企業を検出する新しいネットワーク指標を提案することで、この課題に取り組む。

生物界のエコシステムの複雑な適応システムに着目した Martin と Sunley は、進化生態学のモデルを用いて産業クラスターの進化の概念を検討し、クラスターのライフサイクルモデルで提示されてきた誕生、成長、成熟、衰退、消滅といった段階は、一方向に進むのではなく、八の字の軌跡を描く複雑な適応プロセスであるとした [132]。しかし、クラスター進化を理解するために提案されたこの適応サイクルモデルは、概念として提示されたものであり、現実のデータへの適用可能性やマイクロレベル・マクロレベルの変動過程に関する詳細な実証研究は今後の課題とされている。本論文第5章では、異なる役割を持つ各企業の行動がどのようなクラスター進化につながるかについて定量的分析を行うことで、この課題に取り組む。

2.2 生物と経済の接合過程に見る持続的社會モデルへの試案

前節では、異質な存在とつながることのできる環境が、付加価値を生むイノベーション活動において重要であることを示す先行研究とその課題について論じてきた。本節では、本論文の着想の原点となった考え方とアプローチについて述べる。

2.2.1 生物界の新陳代謝

(1) 極相林のギャップダイナミクス

植物群集は生育するにつれて生育地の環境条件に影響を与え、変化した環境条件に適した植生に遷移していく。古典的生態学によると、光や水、大気、土、温度などの無機的ストレスと資源をめぐる種間競争が遷移の基本的メカニズムであり、その最終段階ではその地点における環境条件に最適な樹木が優占種（典型的には、暗い環境でも定着が可能な陰樹）となり、それ以降樹木種の構成がほぼ変化しない状態にいたる。この状態を極相という [133]。

古典的な極相の概念がある種スタティックな極限状態を表していたのに対し、近年の生態学ではよりダイナミックな極相の状態が論じられている。森林に欠損が生じたとき、すなわち枯死や倒木、落雷等による木の損傷によって森の天蓋部分（林冠）に穴（ギャップ）が生じ、そこに直射日光が差し込み温度が上昇するなど環境条件が変化すると、新しい植物種が参入する余地が生じる。実際にはギャップが生じるタイムスパンが、樹木の生存期間よりも短い場合も多く、極相に達するまでの過程の状態を常に繰り返す森林も珍しくない。近年では、森林生態系に中規模の攪乱をもたらすギャップの役割が、森林の維持機構として特別に重視されている [134]。植生遷移の最終段階である極相だとしても、それ以上変化しないものとするのではなく、むしろ動的な平衡状態にあると理解されており、ギャップを埋めることで得られる新たな均衡状態は、元の極相林の完全な再現にはならない場合も多いことがわかっている（図 2.1）。ギャップ説とは、たえざるギャップの形成と閉鎖を通じて種多様度が保持されるという考えであり、こうした森林動態のあり方を「ギャップダイナミクス (gap

dynamics)」という [1]. 極相林は部分的に壊れては変化することを繰り返しながら、平衡状態付近を行き来しているのである.

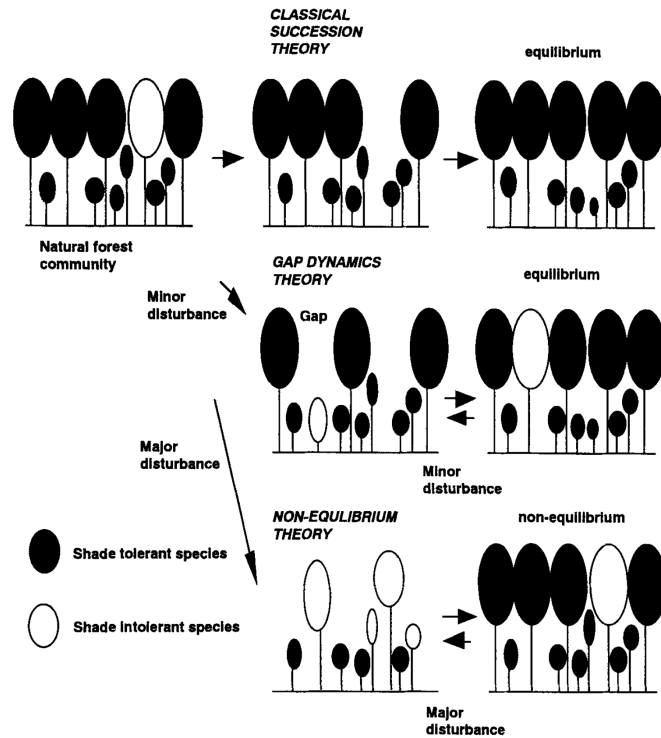


図 2.1: 極相の古典的な遷移, ギャップダイナミクス, 非均衡理論の概念図 [1]

2.2.2 産業界の新陳代謝

(1) 産業の新陳代謝が牽引した日本の産業発展

日本経済は 19 世紀末以降, バブル崩壊に伴う金融危機が深刻化する 1980 年代終わりまで, 約 100 年にわたって高い成長率を維持してきた. この経済成長は同質的な企業や産業の規模拡大によるものではなく, 「産業の新陳代謝」ともいえる, 主導産業の交代をはじめとする大きな産業構造の変化を伴うものであった. たとえば, 第二次世界大戦前, 日本の経済成長を支えた主導産業は, 商業・サービス業から鉱工業に交代し, 第二次世界大戦後は牽引部門の大幅な交代を伴いながら製造業が鉱工業に取って代わった. 産業構造の変化は産業間だけではなく産業内部にも見られ, 19 世紀末から 20 世紀初頭の綿紡績業では 8 割以上の企業が入替わったこと, 主に新規企業の参入により, 同産業の生産性が上昇したことが知られている [135, 12]. すなわち, 産業内外の新陳代謝が日本の経済成長の大きな牽引力となってきたのである [13].

産業毎の推移動向を生物学的観点で捉えた研究として, 産業のクロックスピードの概念がある. ショウジョウバエはヒトを含めた脊椎動物と共通する遺伝子と形態形成の仕組みを持

ち、世代間隔が約2週間と短く進化速度が速いため、生物の普遍的原理を探るモデル生物として頻繁に用いられる。Fineはこのショウジョウバエの進化速度から着想を得て、企業が市場環境の変化に適応して行う技術進歩の速さ（クロックスピード）と結びつけた企業戦略を論じた [20]。クロックスピードは産業毎に大きく異なるが、企業が持つどんな優位も一時的なものに過ぎず、産業の主導権はサプライチェーンの中でクロックスピードに応じた変化にしたがって常に移り変わるとした。クロックスピードは産業の新陳代謝を、産業毎の進化速度の違いと結びつけて捉える概念と言える。

破壊的イノベーションを提唱した Christensen は、なぜ優良企業が失敗するのかという問いに答えるための最初の観察対象として、ディスク・ドライブ業界を選択している。その理由は「ディスク・ドライブ・メーカーは、産業界で最もショウジョウバエに近い存在だ」とする友人からの助言によるものであった [21]。新しい世代が次々に現れる、業界の進化速度の速さに着目したのである。新陳代謝度の高い産業は、予測不可能な発展を遂げる破壊的イノベーションを検証する、モデル産業としての利点があったことが伺える。

(2) 産業の新陳代謝とは

産業の新陳代謝 (industrial metabolism) という概念はもともと、生物が生命を維持するために不可欠な機構である新陳代謝と経済システムの制御機構とのアナロジーに着想を得て提唱されたものである。提唱者の Ayres は 1960 年代に経済圏、産業、そして企業の各レベルにおいて、新陳代謝の定義を試みている。物質やエネルギーを消費することで維持される散逸構造としての生命体における新陳代謝と類似する機構を、経済活動の各レベルに見出したのである。すなわち、物質の流れを伴うエコシステムとして経済圏、エネルギー効率化など持続可能な発展に向けた産業再編、そして原材料や燃料を用いて製品や廃棄物を生じる企業の製造過程に見られる新陳代謝である [22]。ここでの主眼は、経済成長の牽引力としての主導産業変更ないし世代交代による構造変化というよりは、生物にヒントを得た生命維持機構であり、有限なリソースを効率的に使う生存戦略としてのフロー最適化である。構造変化モデルが直線的経済成長を重視したのに対し、フロー最適化は交代を伴う系が定常的に維持される仕組みに注目した点で、より生命体における新陳代謝の本質を捉えたモデルといえる。ただし、課題はこうした新陳代謝の分析に適したデータ取得の難しさと計測可能な指標の構築であった。Ayres はリサイクルの観点から重金属の消費量に注目したが、これは資源として管理されていた重金属に関するデータが利用しやすかったためではないだろうか。本来、「産業の新陳代謝」を検証するためには、広範囲に渡る多種産業の動向を、ある程度長期間にわたって観測する必要があるが、この概念が提唱された当時は定量的な分析に必要な質の高いパネルデータの入手が困難であったと思われる。

(3) 経済成長の不変性と変異性、不完全な複製

森林生態学におけるギャップダイナミクスの考え方は、市場環境の予期せぬ変動に対する企業間取引ネットワークの寸断や維持、更新を考える上で示唆的である。たとえば、金融危

機や自然災害といった何らかの外的変動により取引関係が攪乱されるとき、それまで維持されていた市場の平衡状態（完全競争市場）が崩れ、新たな遷移段階に移行するための探索の有効性が高まるとする。その際、日々の努力により新たな技術開発や事業展開の可能性を探っていた企業は生存に有利な戦略を見つけやすいだろう。（Schumpeter が「経済発展の理論」で着目したのも、均衡を攪乱する動態的契機であり、攪乱の影響とその結果生じる新しい、経済的均衡状態の研究であった [23, 136].）

村上是新古典派経済学における理想化された市場としての完全競争市場を、生態学的極相と対比できるかもしれないと述べている [137]。ここでの着眼点は市場における貨幣以外の交換価値としての文化子（典型的には技術）の模倣過程である。すなわち、互いの技術を模倣しやすい条件下で、無数の企業がそれぞれのシェア拡大を争う時、規模の経済が有利に働く費用逓減の可能性は減少していき、新たな技術の可能性を見いだせない諸企業は、「極相」に収斂していく可能性が高いだろうと考えた。模倣の可能性の飽和（費用逓増）が異なる企業群の共存状態を可能にしているという完全競争における想定を、極相が成り立つ1つの条件としての遺伝子の不変性と対比したのである。（ただし、村上是文化子を遺伝子型に相当するものと考えており、種ないし生物個体を表す表現型にあたる極相の概念との対比には限界があることを指摘している。）

しかし、現実の市場経済では、完全競争のように理想化された技術の模倣は不可能であり、むしろ各企業が、他社のリソースを完全には模倣できないことによって多様性が生じ、企業や企業群が進化していくと考えられる。村上是動学的な経済成長の分析手法に生物学のアプローチを参照する理由として、経済学と進化論との類似性、特に遺伝子の不変性と変異性の二重構造を挙げ、下記のように述べた「変わらないものが変わっていくというパラドックスが進化なのである。変わるものが変わるのには、単なる変化であって進化ではない。不変性を貫こうとする力とその不変なるものを変える力が絡み合う二重の機制が、進化に他ならない」[137]。ここで、不変性を貫こうとする力として想定されているのは、高度な正確性を持つ遺伝子の複製機能のことである。しかし、その複製が完全に正確ではなく、まれに誤りを生じることが予測不可能な不連続的変異につながる。

ここで重要なのは複製の不完全性の程度である。DNA の複製を担う DNA ポリメラーゼには複製の誤りを正す自己修復機能があり、ヒトの場合、最終的に誤りは細胞内の DNA を構成する約 30 億塩基対のうち 1~2 個に抑えられている。ヒトの細胞数はおよそ 37 兆個とされているが [138]、このうち約 1 兆個の細胞が日々入れ替わっているため、毎日 1 兆個は間違えているという試算になる。こうした複製エラーの蓄積は突然変異につながることもあり、生物の多様性を生み出している。

遺伝子、細胞、個体、個体群と様々なレベルにおいて不完全な複製を成り立たせる仕組みは組み込まれており、それが生物の柔軟性や多様性を可能にしている。広い意味ではこれまでにない情報に入れ替わるという点で、新陳代謝機構と言い換えてもよいだろう。では、こうした仕組みは企業や企業間の相互作用においても見られるだろうか。先述のように、組織学習には探索と深化 [80, 75]、企業間ネットワークでは近距離と遠距離 [40, 139]、強いつながりと弱いつながり [113, 80, 98] といった 2 つの拮抗する相互関係のバランスが、イノ

バージョンを可能にする上で大切であることが指摘されている [114]。詳細は本論文の第3章で述べるが、企業間の取引にも不変性と変異性の二重構造は存在する。企業が持つ取引関係の大半は維持されるが、その中の一部の取引を企業は時々入れ替えるからである。これまで見てきたように、不完全な複製は完全な複製を前提として成り立っており、拮抗する両者の作用の絶妙なバランスが進化や形態形成においては重要な役割を果たす。では、企業間取引における複製の不完全さないし新陳代謝度として、最適な割合は存在するだろうか。第3章では、企業間取引データを用いてこの問いを追求し、企業の新陳代謝と存続年数についての仮説を検証する。

極相が崩れる場合として、村上は文化子そのものが創造的破壊によって進化することを上げている。しかし前節で振り返った近年の生態学が論ずる極相におけるギャップダイナミクスの議論を踏まえれば、極相そのものも動的平衡状態にあるのであり、それは新たな遺伝子の出現や進化というよりは、ギャップの出現によるものである。種のレベルで森林の多様性を考える際には、ギャップによる新しい植物種の参入を想定する方が自然であるし、遺伝子の変異が生物個体や種の進化に現れるまでにかかる過程と膨大な時間スケールを考慮するより現実に近い。そこで本論文では、ギャップダイナミクスの考え方に基づいて企業間取引の動態を捉える指標を作成した。その一つの着眼点は、極相林の平衡状態が攪乱の規模を表すギャップの大きさに依存するというものであり、これを企業の新陳代謝に当てはめて考えれば、新陳代謝度の大きさが、変動する市場環境への対応力を示唆する企業の存続年数とどのように関係しているかという問いになる。もう一つの着眼点は、極相林の平衡を乱すギャップという構造が、新しい植物種の参入を促し森林の多様性維持の鍵となっているというものであり、これを企業間ネットワークに当てはめて考えれば、ネットワークの穴である構造的空隙を埋め、地域クラスターやビジネスエコシステムに多様性をもたらす企業の役割は何かという問いになる。本論文ではこれら2つの問いに対し、ネットワーク科学の手法を用いた解析と検証を行う。

2.3 ネットワーク解析の分析方法

ネットワーク科学の観点では、ギャップはネットワークの構造的空隙 (structural hole) [45] と捉えることができる。本節では企業間ネットワークの動的推移に関する研究の課題を述べるとともに、構造的空隙を検出するネットワーク指標について説明し、知識交換を伴う企業間相互作用の分析に必要な、動的推移と静的構造、および両者を踏まえたコミュニティ進化を分析する手法について概説する。

2.3.1 動的推移解析：環境適応と影響予測

(1) ネットワークの動態の定義と測定の課題

地域経済のエコシステムにおける企業間取引ネットワークの柔軟性や鋭敏性が、環境への適応力といった企業のパフォーマンスに与える影響については多くの研究があるが [140,

141, 4, 7], 近年の関連文献調査では, 理論的研究や実証分析の研究蓄積に比して, 大規模データに基づいてマイクロレベルの動態を定量的に調べた研究は少ないことが明らかになっている [8, 10, 9, 142, 143].

Laaksonen はダイナミック・ケイパビリティに関する論文の文献調査を行い, 概念的課題を明らかにする研究の成果に比べて, 実証的課題を論ずる研究はあまり注目されてこなかったことを指摘し, 研究者が実証研究に必要な定量的測定手法に光を当てることによって理論的な曖昧さを克服することが重要であると述べている [142]. Brozovic は, 変化に対応する能力と定義される戦略的柔軟性に関する文献を調査し, 貢献の相対的な希少性と概念的な研究の隆盛から, 実証研究の大幅な増加, 定量的研究の優位性に至るまでの分野の進歩を明らかにした [143]. 同様の傾向は近年のサプライチェーンに関する文献調査でも見られ, 大半の研究が定性的で, 概念的・理論的な発展に焦点を当てたものであること, 検証可能な測定に基づいた研究が不足していることが指摘されている [8, 10, 9]. Tukamuhabwa らは, サプライチェーンのレジリエンス (supply chain resilience: SCRES) を向上させるための 91 の提案を分析し, そのほとんどが概念的・理論的な研究に基づいていることを明らかにした. また, 少数の実証研究のうち, 縦断的なアプローチを採用しているのは 2 つだけであることも明らかにしている [9]. Hohenstein らは, SCRES 研究の文献レビューにおいて, 明確な定義を持つ測定指標に基づく研究の必要性を明らかにしている [10]. 企業間ネットワークのダイナミクス分析に関してはこうした定量分析や指標設計上の課題に加え, グラフ理論に基づく複雑ネットワークの分野が出現した当時は, 大規模な動的データの入手がそもそも困難であったという点も挙げられている [144].

(2) 大規模データを用いた企業間ネットワークの定量分析

近年日本では質の高い大規模な企業データが入手可能になっており, 既存のネットワーク指標や数値シミュレーションモデルを用いた企業間ネットワークの定量的な分析が行われている [55, 145, 146, 147, 148, 149]. このうち, 生物学的モデルにヒントを得た研究としては, 細菌の代謝ネットワークの構造情報を用いて設計されたノード指標とノード分類手法 [47] を企業間取引ネットワークの企業評価に適用し, 異なるクラスターへのリンクが多い企業が, 地域ネットワークのスマールワールド性を向上させる能力があることを示した研究がある [55]. また, 全域木とクラスター内外へのリンク分布から, エッジを骨, 脂肪, 軟骨, 筋肉といった解剖学的組織単位に分類することで, 企業間取引ネットワークの堅牢性への影響を評価している研究もある [148]. しかし, これらの研究はある一時点における企業間取引の静的 (static) なネットワーク構造を解析したものであり, 企業の環境適応や経済活動のプロセスが地域産業の発展にどうつながるかというダイナミクスを捉えているわけではない. 大規模データを用いた動的推移分析としては, 他に 40 年間の M&A と企業間取引データを用いた数値シミュレーションもあるが [149], これは次数をはじめとするネットワーク全体のマクロな特徴の推移を企業規模別に捉えたものであり, マクロの特徴とは異なる可能性がある個別企業のマイクロな取引関係の推移は考慮されていない.

本論文第 3 章では, 企業間取引ネットワークにおける企業活動のダイナミクス, 具体的に

は個別の企業のミクロな取引先交代を計測する新たな指標として「新陳代謝度」を提案し、提案指標に基づくネットワークの経時変化と企業や地域産業のパフォーマンスとの関係を考察する。

2.3.2 静的構造解析：ノード指標とコミュニティ構造

ネットワークを構成するノードの重要度を測る代表的な尺度として、中心概念に基づく中心性分析があげられる。ここでは本論文で提案する指標の比較指標として用いる代表的なネットワーク中心性指標、および構造的空隙を検出するノード指標について説明する。

(1) 媒介中心性

中心性の概念をグラフ理論によって定式化した Freeman は、媒介中心性を 2 点間を結ぶ最短経路のうち、対象ノードを通る経路の割合として定義した [150, 151]。定義式は式 (2.1) で表される。媒介中心性はノード除去によりネットワークが分割される度合いに依拠した指標であり、ノードによってもたらされる情報の流れを評価していると言える。なお、エッジの媒介中心性に基づく枝刈り手法は Newman らによるモジュラリティベースのコミュニティ検出に用いられている [152][153]。

$$Bet_i = \sum_{s,t \in V} \frac{\sigma(s,t|i)}{\sigma(s,t)}, \quad (2.1)$$

V はネットワークを構成するノード群を表し、 $\sigma(s,t)$ はノード対 (s,t) をつなぐ最短経路の総数、 $\sigma(s,t|v)$ はこのうちノード s,t 以外のいずれかのノードであるノード i を通る最短経路の数である。ここで $s=t$ の時、 $\sigma(s,t) = 1$ であり、 $v \in s,t$ ならば、 $\sigma(s,t|v) = 0$ である。媒介中心性はノード i の次数に応じて変動する指標であり、ノード i のネットワーク内でのローカルな位置関係を強く反映する。

(2) Katz 中心性

Katz 中心性 [154] は固有ベクトル中心性の一般化であり、式 (2.2) で表される。ノード i の中心性は、ノード i にエッジを持つ全ノードの中心性の総和として表現される。これはノード i に対し、リンクする全てのノードの中心性を遷移させることに相当し、この遷移を繰り返すことで値が収束する。

$$x_i = \alpha \sum_j A_{ij} x_j + \beta, \quad (2.2)$$

A は隣接行列であり、遷移が収束する時、 x_i は最大固有値 λ_{max} を持つ固有ベクトルとなる。パラメーター α と β は正の定数である。 α は $\alpha < \frac{1}{\lambda_{max}}$ を満たし、中心性の値の大きさを

制御する。 β は全ノードに初期値として与えられる小さな値である，固有ベクトル中心性にこの β 項が加わることにより，入次数が 0 のノードだとしても中心性が 0 になることがなくなり，隣接ノードに中心性を遷移させることができる． Katz 中心性では，中心性の高いノード j にエッジを張るノード i は，ノード j の次数に関わらず一律に大きな中心性を持つ（ここが page rank との違いである． page rank ではノード j の次数が大きいほど遷移する中心性が小さくなるようにペナルティが掛けられており，中心性をノード j の次数で割った値が遷移する [155]）． Katz 中心性は直接的な隣接関係というノード周囲のローカルな情報を強く反映する指標であり，コミュニティ間の差異を捉えてはいない．

(3) Burt の拘束度

構造的空隙を提唱した Burt は，構造的空隙そのものを分析対象とすることはできないため，ネットワークにおける構造的空隙の欠如を表す指標（Burt の拘束度）を考案した [46]． Burt の拘束度はノード i に関するネットワーク内の全てのペアのローカルな拘束度を足し合わせた値であり，式 (2.3) で定義される． ローカルな拘束度とは 2 つのノード間の相互情報量と，両ノードとその隣接ノードとの相互情報量の積和を足した値の二乗である．

$$B_i = \sum_{j \in V_i, j \neq i} (p_{ij} + \sum_{q \in V_i, q \neq i, j} p_{iq} p_{qj})^2, \quad (2.3)$$

$$p_{ij} = \frac{a_{ij} + a_{ji}}{\sum_{k \in V_i, k \neq i} (a_{ik} + a_{ki})}, \quad (2.4)$$

V_i はノード i の隣接ノード集合であり， a_{ij} はノード i からノード j への関係の強さを， p_{ij} は i, j 間の関係量を i の持つ全関係量の総和で除した相互情報量を表す． 重みなし無向グラフの場合， p_{ij} は ij 間にリンクがあれば $1/(i$ の次数)，なければ 0 となる． Burt の拘束度は対象となるノードの次数と共通隣接ノード数を反映し， 1 パスおよび 2 パス先までのノード周囲情報に基づいて計算される． 3 パスより先の情報やコミュニティ差異は反映されないため，大規模な複雑ネットワークにおける構造的空隙を埋め，異質なコミュニティ間をつなぐノードを検出する指標とは言い難い．

(4) ハブとコネクター（Z 値と P 値）

ネットワーク内のノードやリンクが多い場合，ネットワーク構造は複雑になる．そこで類似した役割を持つノードを同じブロックにまとめ，ノード間の関係を単純化して把握する技術として考案されたのがブロックモデルである．ブロックモデルでは，ある 2 つのノードが共通にリンクを持つノード（共通隣接ノード）があれば構造同値であるとし，これに基づくノード間の相関パターンにより，ネットワーク内のノードを分割（ブロック化）する [156]．構造同値のノードはネットワークの結合関係から見た代替可能なノードであると捉えられ，構造同値のノード間に直接結合を持つかどうかは問わない．ブロックモデルはネットワーク

内のミクロな関係構造を縮約する有効な手法であるが、ノード間の直接結合に基づいて計算されるメゾレベルのネットワーク特徴であるコミュニティの構造情報は捉えられない [47].

ブロックモデルが捕捉できないノードの役割に着目し、コミュニティへのリンクパターンに基づくネットワーク構造情報をノード評価に取り入れたのがコミュニティ内リンク度を表す指標である Z 値 (within-module degree Z) と、コミュニティ間リンク度を表す P 値 (participation coefficient P) である [47, 157]. 図 2.2 に Z 値と P 値を用いてノードを評価したネットワークの一例を示す. 各指標での評価値が高いノードほど色を濃く大きく表示している. Z 値の高いノードは所属コミュニティ内にリンクを多く持ついわゆるハブであり, P 値の高いノードはコミュニティ間をつなぐコネクターである.

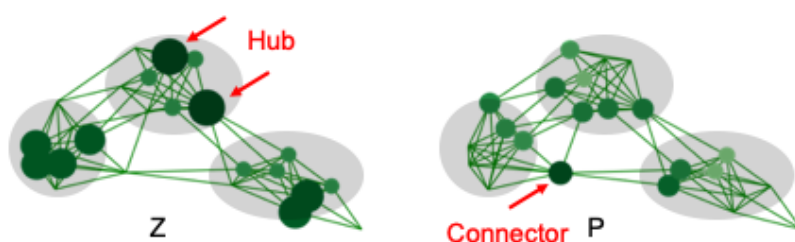


図 2.2: コミュニティ内外のリンク度を測定するノード指標 Z 値と P 値

Burt の拘束度をはじめ, ノードの重要性を測るネットワーク指標の多くが主に隣接ノードを介したノード周囲のローカルな情報, つまり, ミクロレベルのネットワークの構造特徴に基づいて計算されるのに対し, Z 値と P 値はネットワークのメゾレベルでの特徴であるコミュニティ構造をノード評価に活用した指標である. しかし, 両指標が識別するのはノードの持つリンクがコミュニティの中にあるか外にあるかの度合いであり, 複数のコミュニティ間の関係性については捉えられない. 詳細は第 4 章で説明するが, Z 値は定義上 1 つのコミュニティ内のリンクしか評価しておらず, P 値では密集したコミュニティ間のリンクも, 遠く離れて存在しているコミュニティ間のリンクも同等に扱う. このため, 第 1 章で例示したように, 離れたコミュニティをつなぐノードや稀少性の高い共通の友人を捉えることができないという限界がある.

本論文第 4 章では, Z 値, P 値および Burt の拘束度のコミュニティ構造識別力に関する限界を示し, コミュニティに稀少な情報をもたらすノードを捉える新たな指標 PW を提案する. 図 2.3 に本指標の概念図を示す. ノード i が持つ 2 本のコミュニティ外リンクのうち, 指標 PW がより高く評価するのは太い赤線で示された, 遠い関係にあるコミュニティ間のリンクである. 本論文のアイディアは, コミュニティ間の距離をノードのコミュニティへのリンクパターンと掛け合わせることで, ノード評価に活用することである.

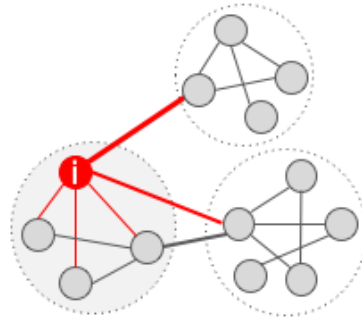


図 2.3: コミュニティ差異を反映したノード指標 PW の概念図

2.3.3 コミュニティ進化の動的モデル

(1) スナップショットとテンポラルネットワーク

Dynamic Community Discovery (DCD) は、ネットワーク科学におけるコミュニティの局所的なトポロジーや変異を追跡する挑戦的課題の一つである。Rossetti と Cazabet は近年 DCD に関する文献サーベイを行い、ネットワーク表現には静的 (static), 重み付きエッジ (edge weighted), スナップショット (snapshot), テンポラルネットワーク (temporal network) の 4 種類があるとした上で、ネットワークの時間変動を扱えるスナップショットとテンポラルネットワークに着目し、概念整理と定式化を行っている [144]。テンポラルネットワーク G は、時間 T におけるノード集合 V とエッジ集合 E のセットとして、式 (2.5) で定義される。

$$\begin{aligned}
 G &= (V, E, T), \\
 V &= (v, t_s, t_e), \\
 E &= (u, v, t_s, t_e), \\
 &\text{with } u, v \in V, t_s \leq t_e, t_s, t_e \in T
 \end{aligned} \tag{2.5}$$

ここで、ノード u とノード v は時間 t_s を始まり、時間 t_e を終わりとするテンポラルノードであり、いずれもノード集合 V に属する。エッジ集合 E はノード u とノード v を結ぶエッジである。また、 t_s と t_e は時間集合 T に属している。テンポラルネットワークは、電子メールや友人関係といった、柔軟で精密な時間情報が得られるネットワークの分析に適している。スナップショットの場合は、ある一時点 t のグラフ $G_t = (V_t, E_t)$ の集合としてネットワークを捉える。スナップショットのアプローチはコミュニティ検出とマッチングの 2 段階に分けられ、ダイナミックコミュニティは、隣接するタイムステップで検出されたコミュニティの特徴をマッチングさせることで定義される。t 時点でのコミュニティは、t 時点でのネットワークのトポロジーに対して最適であると考えられるアプローチである。多重ネットワーク (multiplex network) はスナップショットから時系列のオーダーをなくしたものに相当する。スナップショットは週次/月次/年次といった規則的な時間間隔でネットワークの変動

を捉えるのに適しており、それぞれの時点でのネットワークに静的ネットワークの分析アルゴリズムや計測手法を適用できるという利点がある [144].

(2) コミュニティ進化とノードポジション価値

DCD のフレームワークで用いられるコミュニティの概念を一般化すると、コミュニティとは、他のノードと密に連結（あるいは強く相関）したエッジ集合を共有するノード集合と定義される、ネットワークのメソレベルの構造である（コミュニティ検出はグループ発見、あるいはノードクラスタリングとも呼ばれるが、ここでは統一性のためコミュニティの呼び名を使うことにし、数式表記も C で統一する）。DCD の目的は、ダイナミックネットワークの中の全てのダイナミックコミュニティ DC を識別することである。 DC はノード v と期間 P のペアとして下記の式 (2.6) で定義される。

$$\begin{aligned} DC &= \{(v_1, P_1), (v_2, P_2), \dots, (v_n, P_n)\}, \\ P_n &= ((t_{s_0}, t_{e_0}), (t_{s_1}, t_{e_1}), \dots, (t_{s_N}, t_{e_N})), \\ &\text{with } t_{s_*} \leq t_{e_*} \end{aligned} \quad (2.6)$$

こうした定式化の目的の一つは、コミュニティのライフサイクルを説明する進化イベントを識別することである [144]. Palla らは、コミュニティの誕生、死、成長、収縮、合併、分裂という 6 つの進化イベントでコミュニティ変容を分類することを最初に提案した [158]. こうした進化イベントにはさまざまな識別手法が提案されているが [158, 159, 160, 161], その時々のコミュニティサイズや類似度の閾値によって定義されることが多い。たとえば、Bródka らによる手法 GED (Group Evolution Discovery)[159] では、コミュニティ進化を評価するために、式 (2.7) で定義される包摂性指標 I (inclusion measure) が提案されている。

$$I(C_i^t, C_j^{t+1}) = \frac{|C_i^t \cap C_j^{t+1}|}{|C_i^t|} \cdot \frac{\sum_{v \in (C_i^t \cap C_j^{t+1})} NP_{C_i^t}(v)}{\sum_{v \in C_i^t} NP_{C_i^t}(v)} \quad (2.7)$$

$$\begin{aligned} NP_{n+1}(v) &= (1 - \epsilon) + \epsilon \cdot \sum_{u \in V} NP_n(u) \cdot F(u \rightarrow v), \\ &\text{with } NP_0(v) = 1, \quad u, v \in V, \quad \epsilon \in [0, 1] \end{aligned} \quad (2.8)$$

GED では時点 t におけるコミュニティ C_i^t と時点 $t+1$ におけるコミュニティ C_j^{t+1} のサイズ（コミュニティ内ノード数）比較と、包摂性指標 I の値（ α, β の二つの閾値）の組み合わせによって、コミュニティイベントが識別される。包摂性指標 I は、二つの項の積の形をとり、第一項はコミュニティの積集合に基づく量的評価 (quantity) を、第二項は NP (node position) から求められる重要度に基づく質的評価 (quality) を表している。 NP は式 (2.8) で表される反復計算によって得られる値であり [162], ϵ は 0 以上 1 以下の値をとる定数、 $F(u \rightarrow v)$ はノード u とノード v をつなぐエッジの重さを表す関数である。 $NP_{C_i^t}(v)$ はノード v のコ

コミュニティ C_i^t におけるノード v のポジション価値（重要度）を表している。ノード v のポジション価値は本来はネットワーク内の全ノードのポジションに依存するが、ここでは直接エッジが張られている隣接ノードであるノード u のポジションに限定した計算が行われている [162]。また、重要度の計算には、 NP 以外にも媒介中心性やページランクなど、コミュニティ内のノードポジションを評価する指標が適用できるとされているが [159]、どの指標を用いるべきかについて、計算に要する時間以外の明確な指針は示されていない [163]。

コミュニティ進化イベントの識別にノードの重要度のような質的評価が用いられることは稀であり、多くの識別手法ではコミュニティ類似度に基づく量的評価のみを採用している [164]。Takaffoli らは、コミュニティ類似度に基づく進化イベントの識別手法 MODEC を新たに提案し [161]、ネットワーク中心性等のノードポジション評価指標は進化イベントの予測のために用いている [165]。その際、特徴量はネットワーク中心性等のノード指標以外にも、エッジ密度やクラスター係数、直前の進化イベントなど、コミュニティレベルの変数が多く用いられている。各特徴量の寄与度は、進化イベントを予測する学習モデルにその特徴量が選択された回数から求められるが、性質の異なる多数の特徴量が同時に用いられているため、GED の場合と同様に、ネットワークの中でどのようなポジション価値を有するノードが、進化イベントに寄与するのか説明することは難しい。

本論文第 5 章では、提案指標 PW を含むノードポジション評価指標のみを用いてコミュニティ進化イベントの予測を行う。その際、MODEC で用いられた基準に従ってコミュニティ継承度を定義し、ノードポジション評価指標の中で、コミュニティ継承度の予測に各指標がどの程度寄与するかを検証する。分析対象に用いる企業間取引ネットワークは、10 年分の年次データであり、タイムステップ数が比較的少ないことから、Palla, Bródka, Takaffoli らと同様にスナップショットによるアプローチを採用する。

2.4 本論文の位置づけ

既存研究と照らし合わせた本論文の位置づけについて述べる。

本論文では、地域経済政策やイノベーション研究の中で用いられてきた、「新陳代謝」「ニッチ」「クラスター進化」といった概念に相当する定量的な指標を作成し、ネットワーク科学の手法を使ってその妥当性を検証する。これらの概念はいずれも生命体や生物界のエコシステムに見られるモデルを参考にして、企業間相互作用のあり方を提起するものである。

「新陳代謝」の概念的課題としては、高度経済成長期に見られた主導産業の交代をはじめとする大きな産業構造の変化を伴う新陳代謝が、スピード重視の直線的経済成長をモデルとしていた点、経済政策へ取り入れた際の定義がその時々によって異なるため、統一の見解が得られにくい点をあげた。また、分析面では、産業や地域クラスターといったマクロレベル、メゾレベルの知見に対し、客観的データに基づくミクロレベルでの新陳代謝の定義と検証が不足している点をあげた。本論文では、企業間取引ネットワークにおける企業のミクロな取引先交代を計測する新たな指標として新陳代謝度を提案し、提案指標に基づくネットワークの経時変化と企業や地域産業のパフォーマンスとの関係を考察する。

「ニッチ」については、エコシステムとしての企業ネットワークの最適化を図る観点から見た企業の役割分担と相互作用への着目が、それまでの地域クラスター研究の課題を乗り越えるものであったことを示しつつ、エコシステムに異質性を取り入れるニッチ創出の条件が考慮されていない点をあげた。本論文では、ニッチに相当するネットワーク科学の概念として、Burtの構造的空隙を取り上げ、構造的空隙を検出する既存指標を複雑ネットワークに適用した際の限界を示すとともに、それを乗り越える新たな指標PWを提案する。

「クラスター進化」については、進化生態学のモデルから着想されたクラスターの適応サイクルモデルが概念として提示されたものであり、現実のデータへの適用や、地域クラスターと企業との相互作用を含むダイナミクスに関する実証研究が不足している点をあげた。本論文ではネットワーク科学で近年試みられているコミュニティ進化イベントの識別手法を企業間取引ネットワークに適用し、提案指標PWを含むノードポジション評価指標のみを用いてコミュニティ進化イベントの予測を行う。

第3章 提案指標1：新陳代謝度（リンク推移を反映したノード指標）

第2章では、企業内部の経営資源論 [88, 89] を拡張した概念として、企業間のつながりそのものを資源と捉えるネットワーク資源論 [91, 92] に言及し、既存取引の強化や新たなつながりの構築による企業間ネットワークのダイナミクスが、イノベーションをもたらす重要な要因の一つになっていることを示した。企業間取引ネットワークは、金銭や商品をやりとりするサプライチェーンとしての機能だけでなく、情報や知識の流路、あるいは信頼関係や協力関係の存在を示すものとして注目されている。本章では、日本企業の関係性を捉えるのに最も網羅的かつ信頼性の高い企業間取引データを用いて、企業間の相互関係のダイナミクスを定量的に解析する新たな指標を提案する。具体的には、企業間取引の入れ替え率に基づいて定義した企業の「新陳代謝度」の概念と計算方法を述べ、東日本大震災前後を含む5年間の東北地方、中部地方、九州地方における企業間取引データを用いた分析結果について述べる [166, 167]。存続年数、被消失取引数は本章で独自に定義した特徴量であり、その着想と計算方法についても説明する。なお、企業の新陳代謝度と存続年数との相関関係については、企業間取引データを10年間分に拡大し、統計的有意差検定を含む詳細な解析を行った。また、取引の変動と企業パフォーマンスの複合的な関係を理解するため、各企業の特徴量として新陳代謝度に加え、存続年数、売上増加率、取引数、被消失取引数を用いた非負値テンソル因子分解を行い、潜在的パターン抽出を行った [168]。

3.1 提案指標の定義と分析手法

3.1.1 企業間取引に基づく新陳代謝の基本概念

本節では本章で定義する新たな概念について説明する。まず、本論文の提案指標である、企業間取引に基づく新陳代謝の基本概念を図 3.1 に示す。中心の円は対象企業、その周囲の円は対象企業の取引先企業である。周囲円の色の変遷は取引先企業が年次毎に入れ替わる様子を示し、中心円と周囲円をつなぐ線の太さは取引持続期間の長さを示している。生命体が自身の構成物を一定程度分解し再構成する新陳代謝によって生命を維持しているように、企業も地域経済エコシステムの中で、取引先の一部を入れ替えつつ存続しており、企業の新陳代謝度には企業毎の所与条件に応じた最適値が存在する可能性がある。ある企業にとっての取引先企業の変動には図 3.1 のように、取引先企業の増加、入れ替え、減少の3パターンが考えられる。

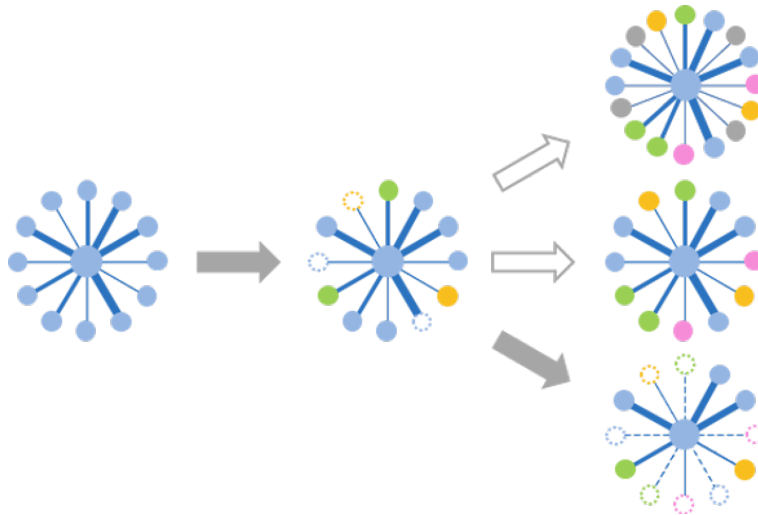


図 3.1: 企業間取引に基づく新陳代謝の概念図

(1) 新陳代謝度 (Metabolism)

新陳代謝度はある企業が設定期間内にどのくらい取引を入れ替えたかを示す指標であり、取引持続度を1から引いた値として以下の式(3.1)で定義される。取引持続度はある企業の企業間取引がどの程度続いたかを表す指標であり、連続する2年における取引先企業類似度の期間平均として算出する。集合の類似度を図る主要な指標には、Jaccard 係数、Shannon 指数、Simpson 係数があり、これらの違いは分母をどう捉えるかである。企業間取引の特徴として、大半の取引関係が翌年も存続する傾向があること、および本論文の着眼点が取引関係がどの程度入れ替わったかであることから、連続する前後年における取引先の和集合を分母に取る類似度として、Jaccard 係数を用いた。 T_y, T_{y+1} はそれぞれ前後年における対象企業の取引先企業を示している。たとえば、データの設定期間が5年の場合、前後年の類似度を計算するため、式(3.1)の n は4となる。企業が取引先を全く変えなければ新陳代謝度は0であり、全ての取引先を入れ替えた場合は新陳代謝度は1になる。また対象期間の前年もしくは後年において企業が取引先を1社も持たない場合、新陳代謝度は計算できない。

$$Metabolism = 1 - \frac{1}{n} \sum_{y=y_1}^{y_n} \frac{T_y \cap T_{y+1}}{T_y \cup T_{y+1}} \quad (3.1)$$

(2) 存続年数 (Duration)

企業の存続年数は通常は倒産や廃業の情報等をもとに把握されるが、本章では今回使用するデータ上で把握できる目安として、各企業が取引先を少なくとも1つ保持している期間を存続年数と定義する。これは対象企業が自社データを申告しなかった年があるとしても、その企業を取引先としてあげている企業が1社でもあれば存続しているとみなすことに相当す

る。逆に、その企業を取引先としてあげる企業が1社もないとしても、自社が取引先を少なくとも1社申告していれば、存続しているとみなす。ある企業が企業が全ての取引先を失うと、取引先企業の類似度の計算結果がNaNを返すため、NaNを返した年数を設定期間の5年から引いた年数が存続年数となる。以降の表では Duration と記載した。

(3) Changer と Holder

上記の定義に基づく存続年数が、本実験の設定期間の最大年数である5年間続いた企業の中で、新陳代謝度が上位5%に属する企業を Changer, 下位5%に属する企業を Holder と定義する。後述する通り、新陳代謝度は山型の分布を示しており、大半の企業は中央値付近の新陳代謝度を持つ。新陳代謝度が極端に高いか低い場合には存続年数は短くなる傾向があるため、Changer と Holder は同水準の新陳代謝度を持つ企業の中では稀な存在である。

(4) 被消失取引数

今回使用する2007年から2016年までの企業間取引データにおいて、ある企業が設定期間中に手放した取引先企業の数(消失取引数)は、ある企業の2016年における取引先企業集合と2007年における取引先企業集合の差の要素数として求められる。これに対し、本章で定義する被消失取引数とは、設定期間中にある企業が失った取引先企業の総数であり、手放した取引先として抽出された企業群に含まれる企業の重複数として求められる。消失取引数がある企業から見た時の評価数であるのに対し、被消失取引数は全企業の取引の取捨選択を踏まえた評価数なので、自社視点を越えたより多くの情報を含んでいる。

3.1.2 群間の統計的有意差検定

(1) 2群間の有意差検定

新陳代謝度が0で取引先を全く替えなかった企業群が、取引先を少なくとも1社は変更する企業群と比較して存続年数に有意差があるかどうか検定を行う。両群のデータは独立で対応関係がなく、正規分布が仮定できない。また、データ数は両群で同じではないため、独立2群間の代表値の差のノンパラメトリック検定である Mann-Whitney の U 検定の(別名 Wilcoxon の順位和検定)を適用した。Mann-Whitney の U 検定の検定統計量は以下の式(3.2)で表される U_1, U_2 のうち、小さい方の値である。

$$\begin{aligned} U_1 &= n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \\ U_2 &= n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \end{aligned} \tag{3.2}$$

ここで n_1, n_2 はそれぞれ第 1 群, 第 2 群のサンプルサイズ, すなわち各群におけるデータ個数を表し, R_1, R_2 はそれぞれ第 1 群, 第 2 群の順位和を表す. 検定統計量が棄却限界値より大きいとき, 帰無仮説は採択され, 検定統計量が棄却限界値以下のとき, 帰無仮説は棄却される.

(2) 等分散検定と分散分析

本章では, 新陳代謝度の分位点水準で分けられた企業群の存続年数について, 群間に有意差があるかどうかを知るため, 分散分析 (analysis of variance; ANOVA) を行う. 分散分析は等分散が前提なので, 等分散を帰無仮説とした検定 (Bartlett 検定) を実施する.

データ総数 N , 群数 (標本数) m の群間と群内の不偏分散の比は, 統計検定量 F 値として式 (3.3) で定義される. 同一母集団から繰り返し標本データを取り出すことで求められる F 値の分布 (F 分布) の両端の面積に対応する p 値が, 群間に差があるかどうかの仮説検定に使われる限界水準である.

$$F = \frac{\text{群間の不偏分散}}{\text{群内の不偏分散}} = \frac{MS_B}{MS_W},$$

$$MS_B = \frac{\sum_{i=1}^m n_i (\bar{X}_i - \bar{X})^2}{m - 1}, \quad (3.3)$$

$$MS_W = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{X}_i)^2}{N - m}$$

ここで群間の不偏分散を表す MS_B (mean squares between) の分子は群間変動であり, 群平均 \bar{X}_i から総平均 \bar{X} を引いて二乗した偏差平方にサンプルサイズ n_i を掛けた値を全ての群について足し合わせて求められる. 分母は群間変動の自由度 $m - 1$ である. これに対し, 群内の不偏分散を表す MS_W (mean squares within) の分子は群内変動であり, データ x_{ij} から群平均 \bar{X}_i を引いて二乗して求められる各群の平方和を足し合わせた値である. 分母は群内変動の自由度であり, データの総数 N から群数 m を引いた値 $N - m$ となる.

(3) 多重比較検定

一元配置分散分析法 (ANOVA) によって群間に差があるかどうかはわかるが, どの群とどの群の平均値に有意差があるかを知るためには, 群間同士に差があるかを別途分析する必要がある. その際, 2 群での検定を何度も繰り返すと, 有意水準が仮定の設定値以上に上昇してしまう多重性の問題が知られている. 本章では, この問題を回避する多重比較検定として代表的な Tukey-Kramer 法を適用する. これは複数の群間の最大の差を考える検定法であり, 群数 m の統計検定量 $max.t$ は計算式 (3.4) で表される.

$$\begin{aligned}
max.t &= \frac{max|\bar{X}_i - \bar{X}_j|}{v\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}, \\
v &= \sqrt{\frac{\sum_{i=1}^m u_i}{\sum_{i=1}^m (n_i - 1)}}, \\
u_i^2 &= \frac{\sum_{k=1}^{n_i} (x_k - \bar{X}_i)^2}{n_i - 1}
\end{aligned} \tag{3.4}$$

ここで u_i^2 はサンプルサイズ n_i , 群平均 \bar{X}_i の i 群における平方和であり, v は m 群分の平方和を足した値を m 群分の自由度の合計で割り, それをルートして得られる, 合算の不偏標準偏差である. $max.t$ の分子は比較する 2 つの標本平均 \bar{X}_i, \bar{X}_j の中で, 最大の差を表している. ある 2 群の差が $max.t$ 値の分布の有意水準 5% に入れば有意差とみなす.

3.1.3 非負値テンソル因子分解のアルゴリズム

非負値テンソル因子分解 (NTF) による特徴パターン抽出方法について説明する. テンソルとは多次元配列のことであり, NTF は多次元配列を低ランクの非負値のテンソル積で近似する分解技術である. 本章では 5 つの企業属性を用いるため 5 次元のテンソルとなるが, ここでは簡略化のため 3 次元配列の場合について説明する. なお, NTF に用いられる行列の要素の値は非負値でなければならないため, 式 (3.5) にしたがって, 5 つの属性データはそれぞれ最小値 0.0, 最大値 1.0 になるよう正規化した.

$$x_{min-max}^i = \frac{x^i - \min(X)}{\max(X) - \min(X)} \tag{3.5}$$

図 3.2 に NTF による分解の概念図を示す. NTF の基底数 R は任意の値であり, データから何種類の頻出相関パターンを抽出したいかを規定するものである. R を増やすことで距離 D を最小化する能力が高まるが, 増やしすぎると結果の解釈が難しくなる. 最適な基底数を求めるため, 本章では近似誤差が最小値付近で飽和する基底数の最小値を選択した.

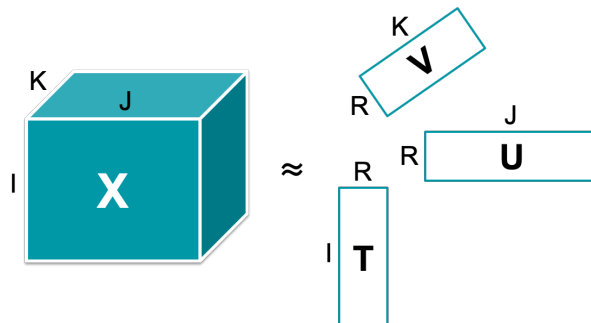


図 3.2: NTF の概念図

$\hat{X} \in \mathbb{R}^{i \times j \times k}$ は推定された因子値を要素を持つ非負値行列であり、 R 次元の潜在変数ベクトルで構成される $T \in \mathbb{R}_+^{I \times R}, U \in \mathbb{R}_+^{J \times R}, V \in \mathbb{R}_+^{K \times R}$ の行列内積の形に分解できる (式 (3.6)). NTF ではこれら T, U, V の非負性を保ったまま式 (3.7) で表される X と \hat{X} の誤差 $D(X|\hat{X})$ が最小になるように T, U, V を学習する. NTF で広く用いられる距離関数 $d(x_{ijk}, \hat{x}_{ijk})$ は, Euclid 距離の 2 乗, 一般化 Kullback-Leibler divergence, Itakura-Saito divergence であるが, 本章で使用する企業属性は数値のばらつきが比較的小さいため距離関数には Euclid 距離の 2 乗を用いる.

$$\begin{aligned}\hat{X} &= [x_{ijk}] = T \otimes U \otimes V \\ &= \sum_r^R t_{ir} u_{jr} v_{kr}\end{aligned}\tag{3.6}$$

$$\begin{aligned}D(X|\hat{X}) &= \sum_i^I \sum_j^J \sum_k^K d(x_{ijk}, \hat{x}_{ijk}) \\ &= \sum_i^I \sum_j^J \sum_k^K (x_{ijk} - \hat{x}_{ijk})^2\end{aligned}\tag{3.7}$$

$$\begin{aligned}\text{minimize}_{T,U,V} \quad & D(X|\hat{X}) \\ \text{subject to} \quad & T, U, V \geq 0\end{aligned}\tag{3.8}$$

距離関数のうち最小化したい部分を目的関数 $F(T, U, V)$ とする. 目的関数を直接最小化することは難しいため, T, U, V はランダムな非負値で初期化し, Multiplicative Update Rule を用いて更新式 (3.9) を繰り返し適用する. 更新式の導出には式 (3.12) に示す補助関数 $F^+(T, U, V, \lambda)$ を導入した. すなわち, t_{ir}, u_{jr}, v_{kr} について補助関数を偏微分して 0 とおき最小化することで更新式を取得した. 補助変数 λ は Jensen の不等式 (3.10) の等式を満たす値であり, 式 (3.13) で定義される.

$$\begin{aligned}t_{ir} &\leftarrow t_{ir} \frac{\sum_j^J \sum_k^K x_{ijk} u_{jr} v_{kr}}{\sum_j^J \sum_k^K \hat{x}_{ijk} u_{jr} v_{kr}} \\ u_{jr} &\leftarrow u_{jr} \frac{\sum_i^I \sum_k^K x_{ijk} t_{ir} v_{kr}}{\sum_i^I \sum_k^K \hat{x}_{ijk} t_{ir} v_{kr}} \\ v_{kr} &\leftarrow v_{kr} \frac{\sum_i^I \sum_j^J x_{ijk} t_{ir} u_{jr}}{\sum_i^I \sum_j^J \hat{x}_{ijk} t_{ir} u_{jr}}\end{aligned}\tag{3.9}$$

$$\sum_i \lambda_i f(x_i) \geq f\left(\sum_i \lambda_i x_i\right)\tag{3.10}$$

$$\begin{aligned}
F(T, U, V) &= \sum_i^I \sum_j^J \sum_k^K \{-2x_{ijk}x_{ijk} + x_{ijk}^2\} \\
&= \sum_i^I \sum_j^J \sum_k^K \left\{ -2x_{ijk} \sum_r^R t_{ir}u_{jr}v_{kr} + \left(\sum_r^R t_{ir}u_{jr}v_{kr} \right)^2 \right\}
\end{aligned} \tag{3.11}$$

$$F^+(T, U, V, \lambda) = \sum_i^I \sum_j^J \sum_k^K \left\{ -2x_{ijk} \sum_r^R t_{ir}u_{jr}v_{kr} + \sum_r^R \frac{(t_{ir}u_{jr}v_{kr})^2}{\lambda} \right\} \tag{3.12}$$

$$\lambda = \frac{t_{ir}u_{jr}v_{kr}}{\sum_r^R t_{ir}u_{jr}v_{kr}} \tag{3.13}$$

3.2 検証実験 1：企業間取引の新陳代謝度の特徴

3.2.1 データセット

本節では株式会社帝国データバンクから提供を受けた、経済規模の比較的近い日本の3つの地方、すなわち、東北地方(青森県, 岩手県, 宮城県, 秋田県, 山形県, 福岡県), 中部地方(愛知県, 富山県, 石川県, 岐阜県, 三重県), 九州地方(福岡県, 佐賀県, 長崎県, 熊本県, 大分県, 宮崎県, 鹿児島県)における、2010年から2014年までの5年間の企業間取引データを用いた。帝国データバンクのデータセットには、申告する企業が取引している主な発注社・受注社を、それぞれ5社まで選択したものが収録されているため、1社につき10社までの取引先が紐付けられている。企業をノード、取引先をエッジとしてネットワークを作成すると、他社から取引先として挙げられた回数が多い企業は多くのエッジを持つことになる。1社あたり10社までという制限により、作成されたネットワークは全ての企業間取引を表すものではなく、ビジネス関係における主要な取引を反映したものとなる。各企業にとって重要な取引先が抽出されてデータベースに登録されていることが知られており、先行研究において、取引先に上げられた企業構成が大企業に偏っている、いわゆる大企業バイアスが生じているわけではないと判断されている [55, 146]。本論文に関連する先行研究に帝国データバンクの企業間取引データを用いた分析は多く、同じデータを用いることで比較検討しやすいという利点がある。

各地域における年度別取引数、発注社数、受注社数を表 3.1 に示す。分析に用いたのは全産業分類を含む重みなしの無向ネットワークである。前処理として、各年における企業間の重複取引、自己ループは削除している。また、本章では各企業による取引の入替率(新陳代謝度)に着目するため、各年において発注社数あるいは受注社数が10社以上の企業に限定して分析を行った。5年間の総取引数は各地方100~200万、企業数は1.7万~3万社程度であり、各年における取引数は20万~40万、企業数は発注社・受注社ともに3000~6000社程度となっている。3地方とも取引数、企業数は増加傾向にあり、東日本大震災が起きた2011

年においても減少はしていないが、特に東北地方において 2011 年から 2012 年の増加率は他の年に比べて低くなっている。

表 3.1: データ概要

Data item	Region	2010-2014	2010	2011	2012	2013	2014
取引数	東北	1,123,164	203,785	224,300	225,619	232,647	236,813
	中部	2,101,354	392,867	417,024	423,002	430,552	437,909
	九州	1,532,390	266,671	305,850	312,686	318,523	328,660
発注社数	東北	16,798	3,219	3,321	3,341	3,429	3,488
	中部	30,429	5,800	5,969	6,089	6,232	6,339
	九州	20,372	3,832	4,035	4,085	4,159	4,261
受注社数	東北	19,681	3,781	3,899	3,918	3,993	4,090
	中部	33,081	6,440	6,544	6,597	6,708	6,792
	九州	25,402	4,851	5,052	5,096	5,149	5,254

(株式会社帝国データバンク提供：時系列比較用企業間取引データ)

3.2.2 企業の代謝度と存続年数との相関

(1) 新陳代謝度の分布

図 3.3 に東北地方における取引先 10 社以上の発注社企業および受注社企業の代謝度の分布を示す。最頻値が 0.18 付近にある山型の分布であり、大半の企業が毎年平均 2 割弱の取引先を変更していることがわかった。また、取引先を全く変えない、代謝度が 0 である企業が一定数おり、その数は 309 社であった (表 3.2)。なお、発注社、受注社それぞれで代謝度の分布はやや異なり、受注社の代謝度の方がやや発注社より低い傾向があったが、2 割弱の範囲に代謝度の最頻値があること、全く取引先を変えない企業が一定数存在するという全体的な傾向は変わらなかった。この傾向は中部地方、九州地方においても同様であり、これらの分布図は補足資料に掲載している (図 A.1)。

(2) 取引先を変えない企業の存続年数

新陳代謝は生物にとっての生命維持機構であり、生命体は新陳代謝を行わずに生存することができないが、企業体にとってはどうか。本章では企業間取引がどれほど入れ替わっているのかを見る指標として代謝度を定義したが、詳細な分析に入る前にまず考えておくべきこととして、ある企業が取引先を全く変えなかった場合、果たして存続することが可能かどうかという観点がある。

そこで、取引先が 10 社以上の企業のうち、設定期間の 5 年間において全く取引先企業を変えていない、すなわち本章で定義する代謝度が 0 となる企業の存続年数を東北・中

部・九州の3地方で比較した（表3.2）。表中の Fixed は全く取引先を変えなかった企業の数，All は全企業数である。東北地方では，取引先を全く変えなかった309社の中で2年存続した企業は235社であるのに対し，5年間存続した企業は僅か9社であり，5年存続率はわずか2.9%であった。全企業7955社のうち，2年存続した企業は997社，5年存続した企業は5461社であり，5年存続率は68.6%であった。固定した取引しかもたない企業と全企業との間に見られた，存続年数と企業数の傾向は，被災地とは離れた場所に位置する中部地方，九州地方においてもほぼ変わらず，固定した取引しか持たない企業の存続率が低いことがわかった。3地域とも全企業で見ると5年存続する企業が最も多いのに対して，固定した取引を持つ企業に限定すると，これとは逆の傾向が見られ，5年存続する企業が最も少なかった。3地域を比較すると，東日本大震災の被災地である東北地方は固定した取引を持つ発注社企業にとって最も存続が厳しい地域であり，九州地方は固定した取引を持つ受注社企業にとって最も存続しやすい地域であるが，全企業との比較からすればいずれの場合も厳しい存続率となっている。

以上の結果より，取引関係の過度な固定化はビジネス上の競争力を失わせる可能性が示された。固定化した取引を持つ企業の5年存続率は特に東北地方において著しく低く，東日本大震災後の激しく変動するビジネス環境に対する適応力が不足していたことが推測される。なお，以上のような企業の取引関係に基づく新陳代謝度と企業の存続率の関係を明らかにした分析は著者が知る限り，初めてである。

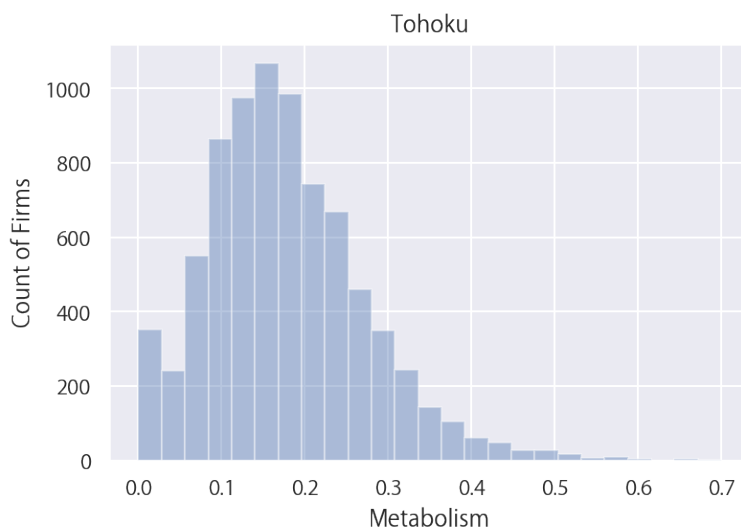


図 3.3: 新陳代謝度の分布

(3) 新陳代謝度と企業の存続年数

企業が1つも取引先を持っていない，あるいは全ての取引先を失った場合，新陳代謝度は計算できず，計算結果がNaNを返すが，このNaNを返す期間の開始年と終了年は，対象企業の取引ネットワーク上での市場参入年と退出年に相当する。たとえば，ある企業の2010

年と 2011 年における取引類似度の計算結果が NaN であり、かつ 2011 年と 2012 年における取引先類似度が NaN ではなかった場合には、その企業は 2010 年には存在していなかったことになる。例として、東北地方の取引先が 10 社以上の企業について、存続年数と新陳代謝度との関係を示す（表 3.3）。ID は各企業に当てた固有番号であり、全ての値は新陳代謝度に基づいてソートしている。簡略のため、大きな傾向が見られた上位 10 社、下位 10 社および ID が 1000 から 1009 までの 10 社の値を表示した。新陳代謝度が低い企業、高い企業の双方において NaN が多く見られ、企業の存続年数が短い傾向があった。

表 3.2: 地方別企業の存続年数と固定化取引

Type	Duration	Tohoku		Chubu		Kyushu	
		Fixed	All	Fixed	All	Fixed	All
Customers	2	103	485	163	793	95	550
	3	22	399	33	694	30	500
	4	8	381	9	573	4	512
	5	2	2400	10	4552	1	2880
Suppliers	2	132	512	165	696	131	539
	3	27	401	41	621	43	527
	4	8	316	17	529	18	445
	5	7	3061	7	5248	14	3966
Survival rate	5	0.029	0.686	0.038	0.715	0.045	0.690
Total count		309	7955	445	13706	336	9919

図 3.4 は、東北・中部・九州地方の新陳代謝度と存続年数の分布を示している。新陳代謝度が計算できる最小値である 2 年しか存続しない企業の新陳代謝度は、3 地方のいずれにおいても、最も分散が大きく、特に新陳代謝度が 0 付近に一定数まとまった企業数が存在していることがわかる。これは前段で取引先を全く変えない企業の存続年数が短くなる傾向と一致している。3,4,5 年を含む全存続年数における新陳代謝度の分布を比較して得られた、3 地方および発注社・受注社のいずれにおいても共通して見られる特徴は、存続年数が長いほど新陳代謝度がある一定の値に収束していく傾向があること、特に発注社企業においては存続年数が長い方が新陳代謝度の中央値が高い傾向があり、その値は受注社企業の新陳代謝度の中央値より高いことである。東北地方に特有の特徴としては、存続年数 2 年における企業の新陳代謝度がより低い値に多く分布していることである。

前段と同様、新陳代謝度が低く、固定化した取引を持つ企業は、東日本大震災後の厳しいビジネス環境に対する適応力が不足していたことが推測される。ここでの新たな知見は、存続年数 2 年の企業の新陳代謝度の分散が大きかったことであり、新陳代謝度が過度に高い企業にも生存期間が短い傾向が見られたことである。取引の過度な変動性 (volatility) は、企業が市場に適切に適応することを妨げる結果につながる可能性がある。すなわち、新陳代謝度は高すぎても低すぎても、市場環境における生存競争に不利に働いていることが示された。

取引先企業数が 10 社以上の企業の新陳代謝度、存続年数、取引数、企業数を東北・中部・

表 3.3: 新陳代謝度と企業の存続年数

ID	2010-2011	2011-2012	2012-2013	2013-2014	Metabolism	Partners	Duration
0	-	-	-	0.2432	0.7568	49	2
1	0.3125	-	-	-	0.6875	20	2
2	-	-	-	0.3333	0.6667	15	2
3	-	-	-	0.3684	0.6316	28	2
4	-	-	0.375	-	0.625	20	2
5	0.3889	-	-	-	0.6111	18	2
6	-	-	0.375	0.4375	0.5937	20	3
7	-	-	-	0.4118	0.5882	20	2
8	-	-	0.3333	0.5	0.5833	20	3
9	-	-	-	0.4286	0.5714	16	2
.....							
1000	0.6316	0.7778	0.7778	0.8235	0.2473	24	5
1001	0.7705	0.7931	0.625	0.8222	0.2473	90	5
1002	0.8438	0.7222	0.6923	-	0.2472	47	4
1003	0.7143	0.8571	0.6875	-	0.247	26	4
1004	0.6667	0.8182	0.7714	0.7561	0.2469	68	5
1005	0.75	0.7292	0.7407	0.7931	0.2467	70	5
1006	0.625	0.7857	0.7692	0.8333	0.2467	20	5
1007	0.9091	0.6429	0.7692	0.6923	0.2466	16	5
1008	0.7403	0.7413	0.766	0.766	0.2466	523	5
1009	0.7989	0.709	0.7654	0.7405	0.2465	254	5
.....							
3285	-	1	-	-	0	10	2
3286	-	-	-	1	0	12	2
3287	-	-	1	-	0	10	2
3288	1	1	-	-	0	10	3
3289	-	-	-	1	0	16	2
3290	1	1	-	-	0	10	3
3291	1	-	-	-	0	14	2
3292	-	-	-	1	0	10	2
3293	-	1	-	-	0	13	2
3294	1	-	-	-	0	13	2

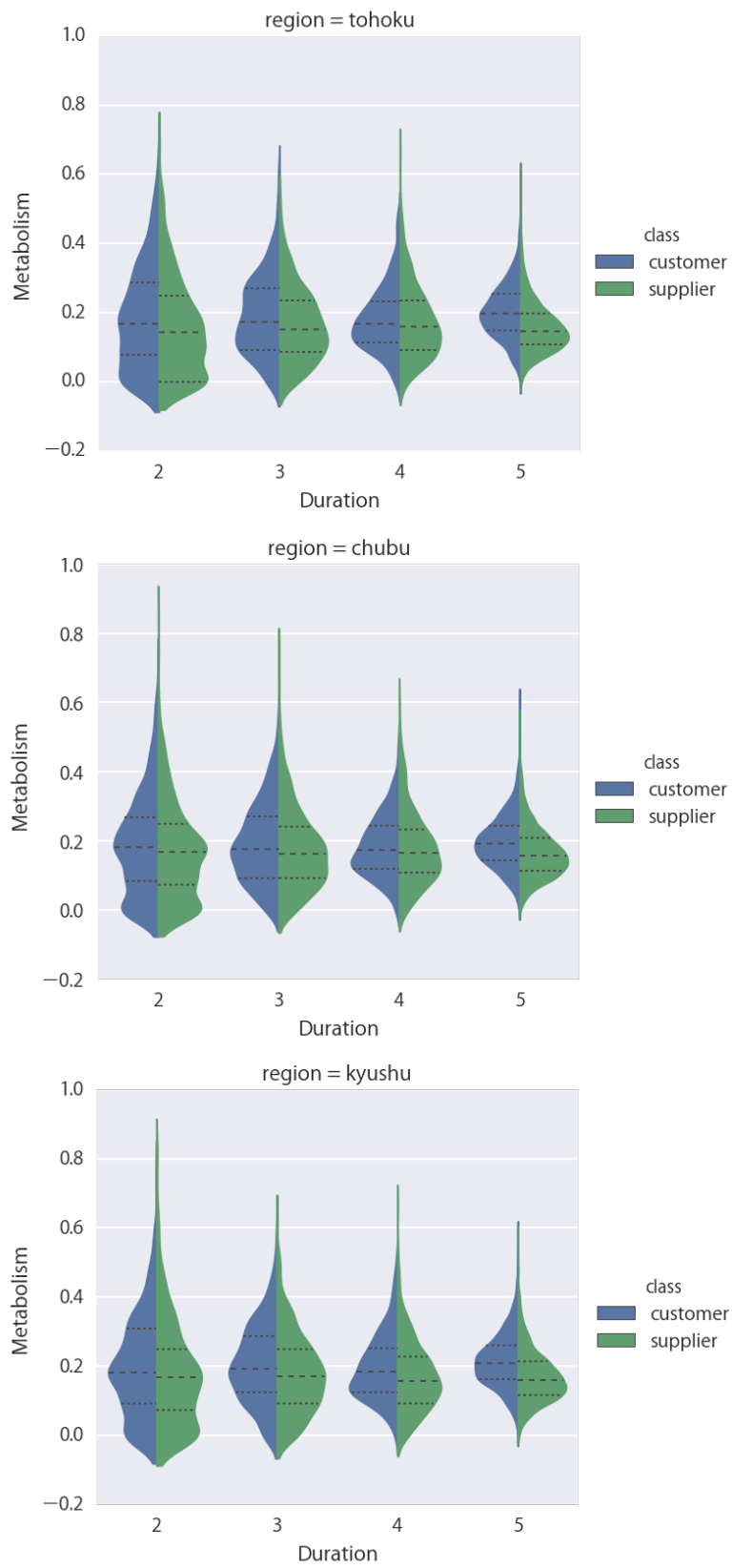


図 3.4: 新陳代謝度と存続年数の分布の地域比較

九州地方の産業別の平均値と比較した（表 3.4）. 3 地方の新陳代謝度に対する全体的傾向は類似しており、いずれも金融保険業、不動産業の新陳代謝度が高く、製造業、卸・小売業の新陳代謝度が低いことがわかった。東北地方の産業における新陳代謝度の最大値・最小値は 3 地方の中でそれぞれ最大・最小であり、産業間の新陳代謝度のばらつきが最も大きかった。また、産業毎に見た場合、新陳代謝度と存続年数の間に相関はなかった。

卸・小売業の新陳代謝度が低く、固定的な取引が多い傾向にあることは、日本における卸・小売業はビジネス上の相互関係をほとんど変化させない企業が多い現実と一致しており、同産業の生産性の低さやイノベーションの遅れにつながっている可能性が示された。一方で、金融・保険業はビジネス関係を頻繁に入れ替えることが知られており、変動する市場への迅速な対応を可能にするビジネス戦略を有していることが示された。

表 3.4: 地域・産業分類別，新陳代謝度と存続年数

Region	Industry	Metabolism	Duration	Transactions	Companies
東北	金融・保険業	0.259	4.347	50.972	67
	不動産業	0.214	4.132	23.053	38
	運輸・通信業	0.215	4.374	35.863	139
	建設業	0.214	4.305	31.867	1221
	サービス業	0.195	4.214	29.991	704
	製造業	0.176	4.336	29.78	1884
	卸・小売業	0.166	4.374	40.809	3536
中部	金融・保険業	0.242	4.426	59.266	133
	不動産業	0.237	4.41	42.943	105
	サービス業	0.207	4.253	31.163	980
	建設業	0.197	4.243	33.821	1757
	運輸・通信業	0.194	4.336	41.641	304
	卸・小売業	0.178	4.419	43.452	5558
	製造業	0.172	4.45	37.964	4560
九州	金融・保険業	0.258	4.378	51.762	94
	不動産業	0.242	4.359	35.312	64
	運輸・通信業	0.212	4.253	37	237
	建設業	0.207	4.165	33.235	1485
	サービス業	0.204	4.316	34.976	980
	製造業	0.183	4.435	34.799	2286
	卸・小売業	0.183	4.413	41.994	4405

(4) 地域毎の新陳代謝度，企業数，存続年数の比較

新陳代謝度は各企業による取引の入替率であるため，対象企業の取引先数はある程度大きい必要がある．本節では取引先企業数が10社以上の企業の新陳代謝度，存続年数，取引数，企業数を比較した．表3.5は，東北・中部・九州地方に位置する18県および3地方それぞれと取引関係を持つ企業を有する主な地域として東京都と大阪府を加えた20地域別の平均値で比較した結果である．産業毎に見た場合と同様，地方別・県別で見た場合にも新陳代謝度と存続年数の間に相関はなかった．各地方とも東京都と大阪府において新陳代謝度が高いことがわかった．東北地方では宮城県，九州地方では福岡県と，いずれも大企業が多く存在する地域の新陳代謝度が高かった．これに対し，中部地方では地域経済をリードしている愛知県の新陳代謝度が低く，取引が固定化していることが示された．愛知県特有の自動車産業を中心とする安定した取引関係の影響が大きいことが推察される．

表 3.5: 地域・県別，新陳代謝度と存続年数

Region	Prefecture	Metabolism	Duration	Transactions	Companies
東北	東京都	0.197	4.429	47.4	1315
	宮城県	0.196	4.24	45.454	1276
	大阪府	0.194	4.314	43.561	431
	岩手県	0.179	4.276	37.856	678
	青森県	0.175	4.353	38.915	904
	福島県	0.167	4.313	33.901	1112
	山形県	0.165	4.352	33.975	875
	秋田県	0.149	4.422	32.982	768
中部	東京都	0.196	4.471	48.742	1691
	大阪府	0.185	4.462	39.823	663
	三重県	0.183	4.316	35.221	1049
	岐阜県	0.18	4.342	35.161	1482
	石川県	0.179	4.386	37.909	984
	愛知県	0.179	4.409	46.103	6016
	富山県	0.167	4.443	42.127	1039
	九州	東京都	0.211	4.459	52.538
大阪府		0.196	4.411	40.468	444
福岡県		0.194	4.326	45.235	3135
大分県		0.187	4.314	38.542	672
宮崎県		0.184	4.312	38.84	665
熊本県		0.183	4.338	47.281	833
長崎県		0.179	4.342	37.535	755
鹿児島県		0.179	4.401	39.691	856
佐賀		0.171	4.35	33.105	562

東北地方においては、秋田県と山形県で新陳代謝度が低かったが、両県は太平洋側の被災地とは逆サイドに位置し、東北地方の中でも東日本大震災の影響は相対的に小さかった地域である。定義式 3.1 より、新陳代謝度が高くなるのは新たな取引先を得た場合、取引先を入れ替えた場合、取引先を失った場合の 3 通りが考えられる。震災後の被災地においては大規模なサプライチェーンの崩壊が起り、取引先を失った企業が多かったため、被災地の新陳代謝度が高くなるのは妥当な結果と言える。これに対し、被災地から離れた地域では被災地に比べて新陳代謝度が低いことから、サプライチェーンの崩壊により取引先を失った企業は一定程度被災地に限定されていたことがわかる。

定義上、サプライチェーン崩壊による取引先喪失は新陳代謝度上昇の直接的要因になる。しかし、被災地から遠くに位置する地域においても新陳代謝度が高い地域が見られたことは、新陳代謝度上昇の要因が取引先喪失という被災地特有の現象によるものだけではないことを示している。中心都市である仙台を中心とした宮城県や東京都、大阪府、福岡県の経済活力が高く、これらの地域で新陳代謝度が高いことを考えると、経済活力の高い各地方の中心地をはじめとする競争の激しいビジネス環境においては、高い水準で取引先を変更していくことが求められていると言える。新陳代謝度は地域経済や地域産業の活性化にも関与している可能性がある。

(5) Changer と Holder の分布と産業別取引数

東北地方の企業の存続年数と新陳代謝度の分布における Changer と Holder の領域を図 3.5 に示す。企業の存続年数は 1~5 までの年数のどれかであるため、企業毎の値をそのままプロットすると分布が年次毎の直線上に並び、新陳代謝度との関係がわかりにくい。このため、図中の青いドットは新陳代謝度でソートした企業順における近傍 20 社の企業の平均値を示している。ドットの分布傾向を見ると、大半の企業は Majority と記載した灰色の枠で囲まれた領域に存在しており、新陳代謝度は 0.1~0.3 程度で、存続年数は長い傾向があることがわかった。新陳代謝度が 0 の企業は存続年数が非常に短く、新陳代謝度が 0.4 以上の企業も存続年数が相対的に短くなっていた。これに対し、赤い枠で囲まれた Changer と Holder に相当する領域にいるのは存続年数が 5 年である企業の中で、新陳代謝度が上位 5% もしくは下位 5% に入る企業であり、同水準の新陳代謝度を持つ企業群の中では稀少な存在となっていることがわかる。

表 3.6 は、Changer と Holder それぞれの新陳代謝度、該当企業数、取引先企業数である。新陳代謝度は Changer が 0.3502 以上、Holder が 0.077 以下であった。また、Changer の取引先企業数は Holder に比べて多い傾向があることがわかった。

表 3.6: Changer と Holder の比較

Class	Metabolism	Companies	Transactions	Duration
Changers	0.3502 - 0.5515	107	18 - 286	5
Holders	0 - 0.077	107	11 - 63	5

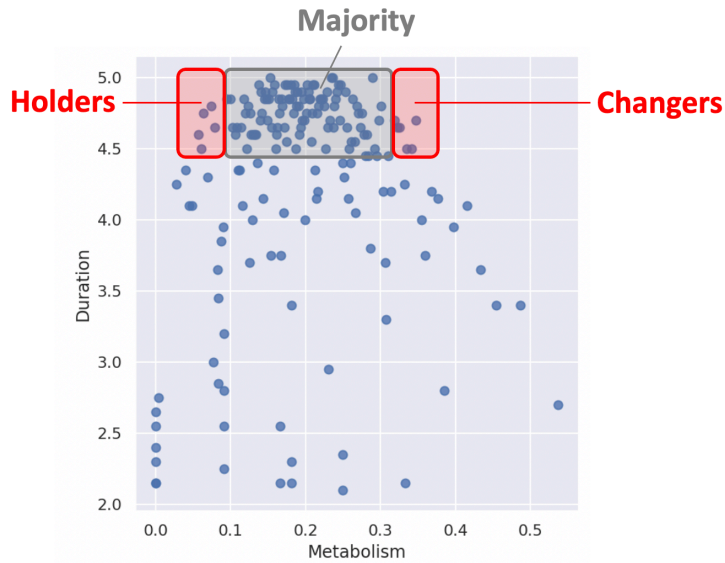


図 3.5: Changer と Holder の存続年数と新陳代謝度

Changer と Holder の取引先企業数を産業分類毎に比較した (図 3.6) . なお, 公務は 2011 年に東北地方に参入した政府主導の復興事業であり存続年数が最大 4 年であるため, 5 年間存続が条件である Changer および Holder は存在しない. 東北地方の企業を産業分類別に比較すると, 建設業は新陳代謝度が高く, Changer が多いことがわかった. 東日本大震災の影響で急増した建設需要が, 建設業界において Changer を生みやすくした可能性がある. その一方, 卸・小売業, 飲食店は新陳代謝度が低く, Changer に比べて Holder がかなり多いことがわかった.

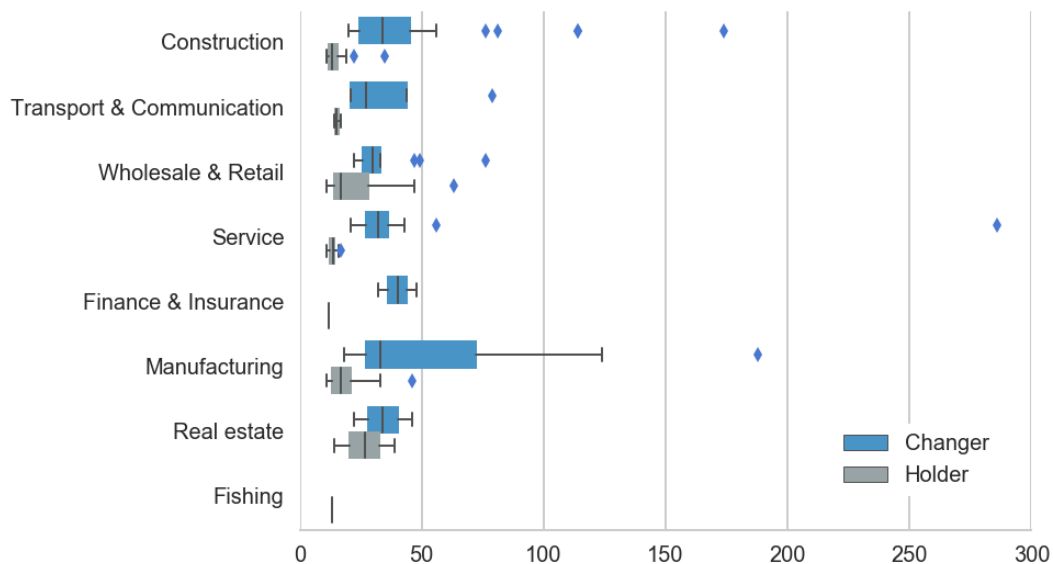


図 3.6: Changer と Holder の産業別取引先企業数分布

この結果は、企業間取引に変化がほとんど見られず、生産性の低さとイノベーションの遅れが認められる日本の卸・小売業・飲食店の実態と一致している。取引数の分布を産業分類毎に比較した結果、Changer は Holder に比べて取引数が総じて多いことがわかった。存続年数が長い企業の中では、取引先を多く持つ企業の方が取引先を入れ替える傾向が強いことが読み取れる。

表 3.7 は東北地方の主要な産業別・県別に見た取引先 10 社以上の企業の平均新陳代謝度、平均存続年数、企業数、および Changer と Holder の数である。データは平均新陳代謝度の高い順に並べており、産業別・県別の企業数に対して Changer もしくは Holder の数が大きい場合には太字で掲載している。Changer 数が多い産業や地域では、平均新陳代謝度が高く、Holder 数が多い産業や地域では平均新陳代謝度が低い傾向が見られた。すなわち、建築業、東京都、宮城県は Changer が多く、卸・小売業、福島県、山形県、秋田県は Holders が多い傾向があった。

表 3.7: Changer と Holder の産業別・県別新陳代謝度と存続年数

Industry	Metabolism	Companies	Changers	Holderes
運輸・通信業	0.215	139	5	3
建設業	0.214	1221	61	24
不動産業	0.214	38	2	2
サービス業	0.195	704	35	24
製造業	0.176	1884	59	63
卸・小売業, 飲食店	0.166	3536	90	153
Prefecture	Metabolism	Companies	Changers	Holderes
東京都	0.197	1315	67	29
宮城県	0.196	1276	67	29
大阪府	0.194	431	16	12
岩手県	0.179	678	16	19
青森県	0.175	904	28	30
福島県	0.167	1112	35	57
山形県	0.165	875	19	34
秋田県	0.149	768	9	46

3.2.3 新陳代謝度と地域経済

前節では、東北・中部・九州地方の県別比較により経済活力が高い地域で新陳代謝度が高い傾向が見られた。そこでこれら 3 地方における合計 18 の県毎に新陳代謝度と売上高の平均値をプロットしたのが図 3.7 である。

県別平均では新陳代謝度と売上高に正の相関が見られ、新陳代謝度が高い県ほど売上高が高い傾向があった。売上高が大きい大企業を多く有する地域が、新事業や新市場の開拓をは

はじめとするビジネス上のイノベーションを起こす可能性がある一方で、売上高の小さい中小企業の多い地域ではイノベーションから取り残されている状態にあることが伺える。

取引先の構成変化は、その企業の経営戦略の変更と関連していると考えられるため、新陳代謝度の高い企業は経済を活性化させるイノベーションを起こす確率が高いことが推察される。実際、東北地方、九州地方における経済の中心地域である宮城県、福岡県が新陳代謝度一売上高座標の最も右上に位置していることは、両県に存在する企業の高い新陳代謝度が、地域の強力な経済活力と適応力に関連していることを示している。一方で、中部地方の経済の中心地である愛知県はこれとは異なる傾向を示し、最も高い売上高を有するが、新陳代謝度が低いため、座標上では外れ値のようにになっている。愛知県がトヨタ自動車株式会社をはじめとする少数の強力な企業が牽引する、製造業中心の特異な経済構造を持っていることを考えると、愛知県の市場環境は競争がより少なく、ビジネス上の相互作用は固定的で安定したものになっていることが推察される。

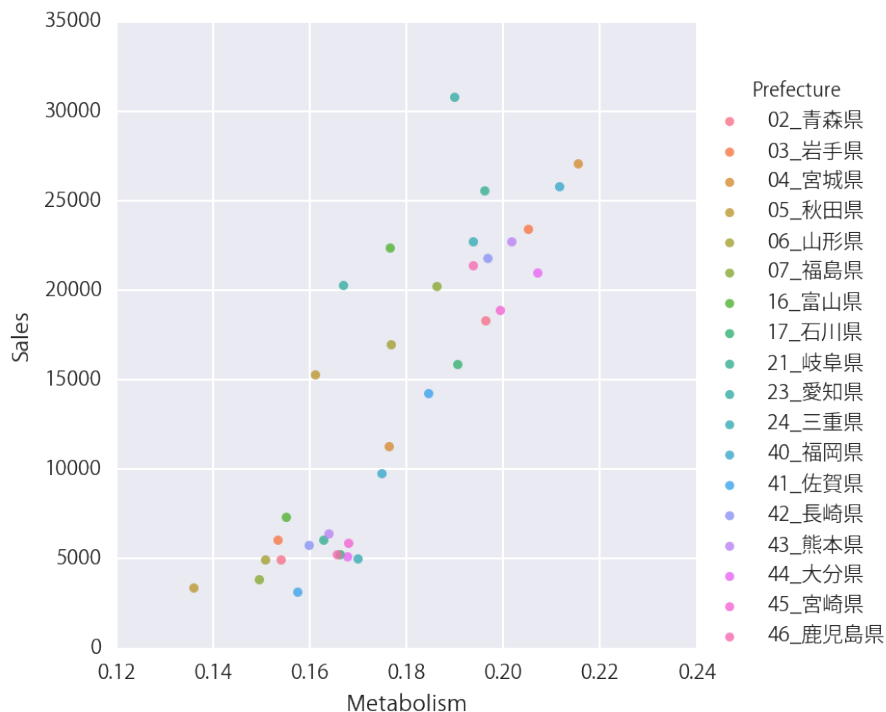


図 3.7: 新陳代謝度と売上高の県別比較

(1) 取引の推移から見た震災復興事業の特徴

東日本大震災から10年間の復興期間（2011年～2020年度）には、31兆円の復興予算が投入された。特に集中復興期間として位置付けられた前半5年間の復旧・復興事業は、安倍政権下で「復興の加速化」を理由に引き上げられ、26兆円を超える規模となった。このような政府による大型の予算投入および事業参入は、地域の取引関係にどのような影響を及ぼ

したろうか。本節では、企業間取引データ上でその規模を確認し、新陳代謝度の観点でどのような特徴が見られたか分析する。

図 3.8 は東北地方の発注社企業の参入取引数と退出取引数の産業別年次推移である。公務 (public sector) は東日本大震災の起きた 2011 年に新しく作られた業種であり、その年の参入取引の中で最も大きな比率を占めている。一方、退出取引は公務以外の全産業において 2012 年が最も多い。政府による多数の復興支援事業関連の参入取引が、震災前から存在していた既存産業の取引に取って代わった可能性がある。また、公務の新陳代謝度は他産業に比べて低く (取引持続度が高く)、取引数は多かった (図 3.9)。政府による東日本大震災の復興事業は大規模であったが、固定的取引が多く、イノベーティブな事業に見られる柔軟なビジネス関係は少なかったことが伺える。

A	481	366	341	420	177	305	220	208
B	50	35	40	44	8	32	23	39
C	48	47	86	53	55	106	56	33
D	90	93	80	83	53	121	84	88
E	10817	11189	11792	12387	10407	11884	9955	11119
F	6447	6538	6591	6372	6205	6570	5996	6028
G	11521	11611	12030	11139	10321	11737	9745	9910
H	358	353	423	357	260	232	239	450
I	601	543	517	653	480	515	430	534
J	1963	2236	2157	2151	1503	2017	1689	2051
K	151	115	139	181	110	167	144	118
L	6900	6144	6147	6101	4201	5043	3981	4464
M	15317	1445	1321	1536	0	1124	1136	1193
	2011_2010	2012_2011	2013_2012	2014_2013	2010_2011	2011_2012	2012_2013	2013_2014

図 3.8: 参入・退出取引の産業別年次推移 (右図: 参入取引, 左図: 退出取引)

A:農業, B:林業・狩猟業, C:漁業, D:鉱業, E:建設業, F:製造業, G:卸・小売業・飲食店, H:金融・保険業, I:不動産業, J:運輸・通信業, K:電気ガス水道業, L:サービス業, M:公務

※カラム名は右図: 参入年_参入前年, 左図: 退出前年_退出年となっている。



図 3.9: 産業別発注社企業

3.3 検証実験 2：新陳代謝度の有意差検定と取引パターン分析

前節までの分析では新陳代謝度の全体的な分布傾向を把握し、企業の存続年数、取引数、地域経済や東日本大震災の復興事業との相関関係の特徴を捉えてきた。本節前半では、地域のビジネスエコシステムを考える上で特に重要な新陳代謝度と存続年数との関係に焦点を絞り、統計的に有意な差があるか比較検定を行う。本節後半では、新陳代謝度を上昇させる3つのパターンを区別する3つの変数を追加し、高次元データの因子分解技術を用いて、企業の取引変動と存続年数の関係を分析する。

3.3.1 データセット

ここまでは、提供を受けたデータにおける企業数が多い帝国データバンクのデータを用いた。以降の研究では株式会社東京商工リサーチ (TSR) の企業間取引データを用いる。TSRのデータセットでは、企業数は帝国データバンクのデータに劣るものの、各企業の取引先は仕入先と販売先、株主に区分されており、申告する企業が取引している主な取引先を、それぞれ20社まで選択したものが収録されている。今回は取引先が株主であるものは対象データから除外し、仕入先・販売先合わせて40社までの取引先が収録されているデータを用いた。TSRのデータは、主な取引先が発注社・受注社毎に各5社まで紐付けられている帝国データバンクの取引データよりも情報量が多く、よりきめの細かな分析が可能となっている。また、今回 TSR から提供を受けたデータの対象期間は2007年～2016年までであり、前節までの研究の2倍の期間である10年間の長期間にわたる解析が可能である。本節では対象

地域を東北地方に限定し、エッジリストのどちらか一方に所在地が東北6県のいずれかである企業が入っていれば分析対象データに含めることとした。今回の分析対象データである各年毎の企業数と取引数を下記に示す。

表 3.8: 東北地方の企業間取引データ

Year	Companies	Transactions
2007	147,328	397,406
2008	147,734	412,653
2009	151,115	415,924
2010	154,390	418,702
2011	155,135	420,544
2012	157,592	424,646
2013	159,758	431,576
2014	159,976	440,914
2015	159,665	451,023
2016	158,757	472,213

(株式会社東京商工リサーチ提供)

3.3.2 存続年数別新陳代謝度の分布

はじめに、新陳代謝度と存続年数の関係について、帝国データバンクのデータ分布と同様の傾向が観察されるかどうかを調べるため、平均取引数が10社以上の企業12,778社に限定して存続年数毎の新陳代謝度の分布を violin plot で描画した。その結果、データの分布傾向は帝国データバンクと同様であり、存続年数が長いほど新陳代謝度の平均値が高く、分散が小さくなる傾向が見られた(図3.10)。なお、新陳代謝度は連続する2年における前後年の取引関係から計算するため、存続年数が1年の企業は描画対象から除外されている。

次に、取引数10社以上の企業を新陳代謝度で15分割し、各値範囲における企業数を存続年数別の棒グラフで描画した。10年存続する企業が大半であるため、描画範囲を2つに分け、10年存続年数が含まれる場合と含まれない場合の2通りで描画した(図3.11, 図3.12)。10年存続した橙色、および9年存続した黄色の棒グラフの企業は新陳代謝度が0.1付近に最頻値があり、存続年数が8年以下の企業は存続年数が短くなるにつれて新陳代謝度が小さくなる傾向が見られた。前節と同様に、長期間存続するために一定程度の新陳代謝度が必要であること、新陳代謝度が低すぎる場合には存続年数が短くなる傾向があることが示唆された。

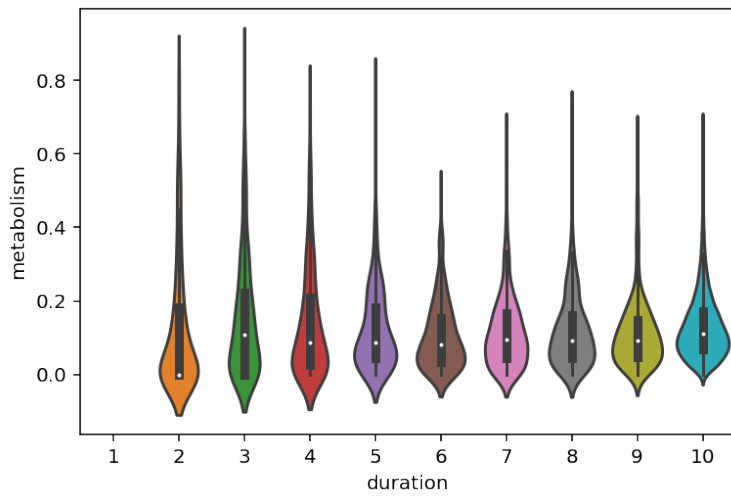


図 3.10: 東北地方の新陳代謝度と存続年数の分布

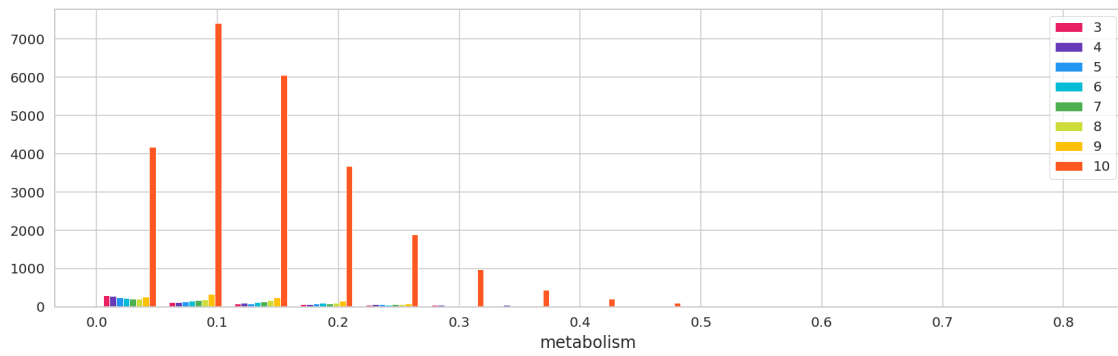


図 3.11: 東北地方の新陳代謝度と存続年数の分布 1

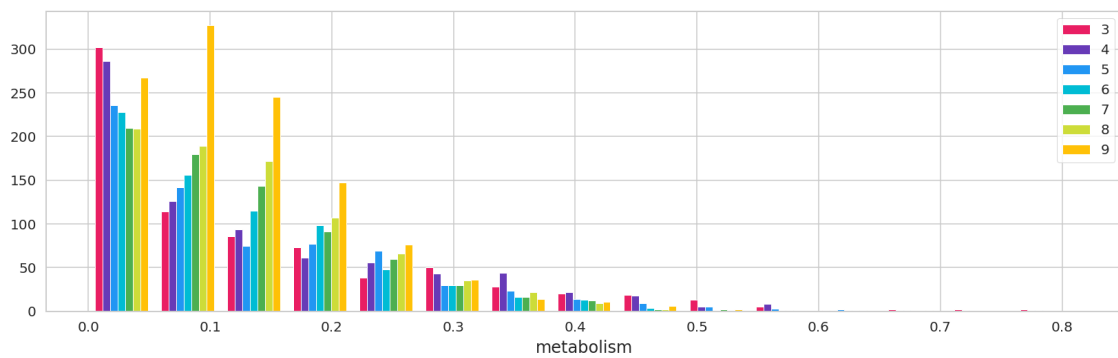


図 3.12: 東北地方の新陳代謝度と存続年数の分布 2

3.3.3 新陳代謝度の統計的有意差検定

(1) 取引先を替えない企業群の統計検定

前節で TSR データにおいても新陳代謝度と存続年数の関係について、帝国データバンクのデータと同様の分布傾向があることが確認できた。以降の分析では、こうした傾向に統計的有意差があるかどうか検定を行う。本節では、新陳代謝度が0で取引先を全く替えなかった企業群と、取引先を少なくとも1社は変更する企業群との独立2群間において存続年数の代表値の有意差検定を行った。対象企業は、10年間の期間中に存在した東北地方の全企業155,883社（東北地方の企業と取引関係のある他地域の企業を含む）のうち、2年以上存続し新陳代謝度が計算できた、取引数10以上の企業10,188社である。なお、取引数10以下の企業も含む存続年数2年以上の企業140,251社の検定結果は補足資料の表A.1に掲載している。

はじめに正規性の仮定を必要としないノンパラメトリック検定である Mann-Whitney の U 検定を行った。ただし、Mann-Whitney の U 検定は等分散が前提とされているため、両群の分散が等しいことを仮定しない t 検定である Welch 検定も平行して実施した。Welch 検定は正規分布を前提としているがある程度の分布の歪みにも対応でき、さまざまな条件下での代表値の比較に対して基本的に安定した高い精度を維持することが知られている [169]。分析の結果、新陳代謝度が0の企業834社と新陳代謝度が0より大きい企業10,009社の2群において、存続年数の代表値はどちらの検定においても $p\text{-value} < 2.2\text{e-}16$ で有意差が認められた (表3.9)。

表 3.9: 独立2群間の有意差検定

条件式	企業数	平均存続年数
$metabolism > 0$	10,009	9.303
$metabolism = 0$	834	4.740
Welch 検定	$t = -33.911$	$p\text{-value} < 2.2\text{e-}16$
Mann-Whitney 検定	$\text{statistic} = 660297$	$p\text{-value} < 2.2\text{e-}16$

(2) 新陳代謝度の分位点別企業群の多群検定

次に、新陳代謝度の値の分位点で企業群を分類し、各群の存続年数に有意差があるかどうかを検定するため、多重比較検定として代表的な Tukey-Kramer 法を適用した。新陳代謝度の値に従ってデータを四分位点および上位下位ともに2.5%、5%、10%の範囲で区切った合計9つの分位点に分割した。分位点毎の基本統計量を表3.10に示す。

今回着目するのは、新陳代謝度が存続年数に与える影響であり、因子は新陳代謝度の1つであるが、一元配置分散分析 (analysis of variance: ANOVA) は等分散が前提とされているため、はじめに Bartlett の等分散検定を行った。その結果、自由度8の対象データにおいて、Bartlett's K-squared = 1003.2, $p\text{-value} < 2.2\text{e-}16$ であり、等分散性は棄却された。

そこで、不等分散かつサイズの異なる多群間の比較を行うノンパラメトリックの一元配置分散分析である、クラスカル・ウォリス法 (Kruskal-Wallis one-way analysis of variance) を適用した。その結果、Kruskal-Wallis chi-squared = 1738.5, df (自由度) = 8, p-value < 2.2e-16 であり、群間の代表値の分布に差はないとする帰無仮説は棄却された。

表 3.10: 新陳代謝度の分位点別企業数と基本統計量

分位点	企業数	平均値	中央値	標準偏差
2.5	510	0.0000327	0	0.000426
5	506	0.0175	0.0185	0.00558
10	1530	0.0435	0.0442	0.00970
25	2547	0.0824	0.0824	0.0131
50	2548	0.135	0.134	0.0184
75	1529	0.202	0.2	0.0219
90	509	0.273	0.270	0.0169
95	255	0.334	0.333	0.0172
97.5	254	0.456	0.430	0.0871
NA	328	NA	NA	NA

(3) 新陳代謝度の分位点別企業群の多重比較検定

Kruskal-Wallis 検定によって群間に有意差があることはわかったが、どの群とどの群の平均値に有意差があるかを知るために、多重比較検定として代表的な Tukey-Kramer 法を適用した。図 3.13 は分位点水準での各企業群における存続年数の分布である。

図の上端のアルファベットの昇順は平均の昇順に対応し、平均間に有意差がない場合には同じアルファベットが付与される。9つの企業群の中で存続年数が最も短いのは新陳代謝度が最も小さい分位点 2.5 の企業群であり、新陳代謝度が最も大きい分位点 97.5, 分位点 95 の企業群がこれに続いている。分位点 2.5 の企業群はほとんどが新陳代謝度=0, すなわち取引先を全く変えなかった企業であり、分位点 97.5 の企業群は平均新陳代謝度が 0.456 であることから、半数ほど取引先を入れ替えた企業である。存続年数が最も長いのは分位点 25 および分位点 50 の企業群であり、両群間に有意差はなかった。

9つの分位点で分けられた企業群のすべての組み合わせにおける統計量を表 3.11 に示す。各分位点での群間で有意差があるものとなないものがあつた。

平均存続年数の群間有意差の 95%信頼区間によるプロットを図 3.14 に示す。今回は 9 群の多重比較検定なので、 $9 \times 8/2 = 36$ 通りの組み合わせの 2 群について、差の推定値を・で、95%信頼区間の下限と上限を (-) で示している。点線は平均の差がゼロであることを意味するため、95%信頼区間が点線を跨がないことは有意水準 5% で有意であることを意味する。例えば、分位点 50 の企業群は分位点 2.5 の企業群に比べて存続年数が 4.5 程度高く信頼区間は狭く、0 を跨いでいないので統計的に有意に平均存続年数が長いと言える。

表 3.11: 新陳代謝度の分位点水準での多重比較検定

Null hypothesis	Estimate	Std. Error	t value	Pr(> t)
5 - 2.5 == 0	4.528	0.117	38.781	< 0.001 ***
10 - 2.5 == 0	4.520	0.095	47.511	< 0.001 ***
25 - 2.5 == 0	4.663	0.090	51.656	< 0.001 ***
50 - 2.5 == 0	4.727	0.090	52.364	< 0.001 ***
75 - 2.5 == 0	4.524	0.095	47.543	< 0.001 ***
90 - 2.5 == 0	4.257	0.117	36.514	< 0.001 ***
95 - 2.5 == 0	3.708	0.143	25.981	< 0.001 ***
97.5 - 2.5 == 0	2.788	0.143	19.513	< 0.001 ***
10 - 5 == 0	-0.008	0.095	-0.079	1.000
25 - 5 == 0	0.135	0.091	1.491	0.841
50 - 5 == 0	0.199	0.091	2.196	0.376
75 - 5 == 0	-0.004	0.095	-0.043	1.000
90 - 5 == 0	-0.271	0.117	-2.319	0.300
95 - 5 == 0	-0.820	0.143	-5.738	< 0.001 ***
97.5 - 5 == 0	-1.739	0.143	-12.156	< 0.001 ***
25 - 10 == 0	0.143	0.060	2.369	0.273
50 - 10 == 0	0.206	0.060	3.430	0.015 *
75 - 10 == 0	0.003	0.067	0.051	1.000
90 - 10 == 0	-0.263	0.095	-2.766	0.110
95 - 10 == 0	-0.812	0.126	-6.455	< 0.001 ***
97.5 - 10 == 0	-1.732	0.126	-13.736	< 0.001 ***
50 - 25 == 0	0.064	0.052	1.224	0.943
75 - 25 == 0	-0.139	0.060	-2.311	0.305
90 - 25 == 0	-0.406	0.090	-4.493	< 0.001 ***
95 - 25 == 0	-0.955	0.122	-7.814	< 0.001 ***
97.5 - 25 == 0	-1.874	0.122	-15.309	< 0.001 ***
75 - 50 == 0	-0.203	0.060	-3.371	0.018 *
90 - 50 == 0	-0.470	0.090	-5.200	< 0.001 ***
95 - 50 == 0	-1.019	0.122	-8.336	< 0.001 ***
97.5 - 50 == 0	-1.938	0.122	-15.830	< 0.001 ***
90 - 75 == 0	-0.267	0.095	-2.802	0.100
95 - 75 == 0	-0.816	0.126	-6.482	< 0.001 ***
97.5 - 75 == 0	-1.735	0.126	-13.763	< 0.001 ***
95 - 90 == 0	-0.549	0.143	-3.846	0.003 **
97.5 - 90 == 0	-1.468	0.143	-10.273	< 0.001 ***
97.5 - 95 == 0	-0.919	0.165	-5.574	< 0.001 ***
Signif. codes	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'

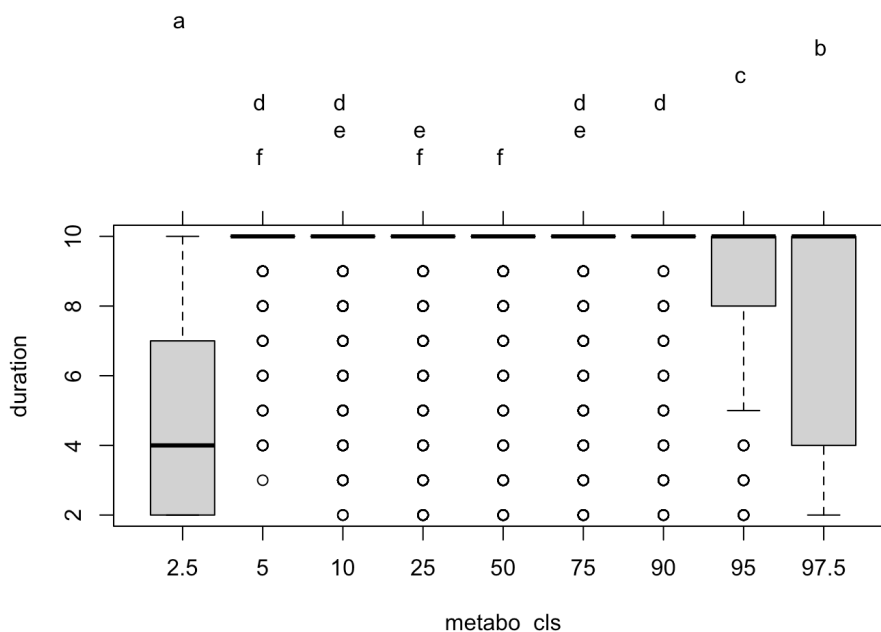


図 3.13: 新陳代謝度の分位点と存続年数の多重比較検定結果

全体的な傾向として、分位点 2.5 の企業群は他の全ての分位点の企業群と有意差があり、平均存続年数が短かった。一方、分位点 2.5 と比べて最も大きな有意差があり、平均存続年数が長い分位点 25, 50 との比較を行ったところ、分位点 25, 50 との有意差は、分位点 90, 95, 97.5 と平均新陳代謝度が大きくなるにつれて広がっていることがわかった。分位点 90 の企業群の平均新陳代謝度は 0.273 であり、分位点 95, 97.5 の企業群はさらに高い平均新陳代謝度を持つことから、取引先を約 3 割以上入れ替える企業群は、新陳代謝度が大きくなるにつれて、存続年数が短くなる傾向があることがわかった。

以上の結果より、存続年数が長い企業群の特徴として、新陳代謝度が小さすぎず大きすぎない中程度の値、すなわち 1 割程度は取引先を入れ替える傾向があることが明らかになった。

Tukey-Kramer 法は多重比較検定で最も多く使われる検定手法ではあるが、正規性と等分散性が仮定されている。これに代わる方法として、分散が等しくないときは Games-Howell 法、ノンパラメトリックである場合は Steel-Dwass 法が推奨されている。これらについても同様に検定を行った結果、Tukey-Kramer 法と大きく変わらない有意差が得られており、詳細は補足資料の表 A.4 および表 A.5 に掲載している。また、変数に企業年齢を加え、同様の多重比較検定を説明変数と目的変数の組み合わせを代えて以下の 3 つのパターンについても Tukey-Kramer 検定を行った。その結果、企業年齢分位点 15 の比較的若い企業群で新陳代謝度が有意に高いことがわかった。分析結果の詳細は下記に示した補足資料に掲載している。

- 目的変数：新陳代謝度 / 説明変数：存続年数（補足資料：表 A.3, 図 A.4, 図 A.5）
- 目的変数：企業年齢 / 説明変数：新陳代謝度分位点（補足資料：図 A.6, 図 A.7）
- 目的変数：新陳代謝度 / 説明変数：企業年齢分位点（補足資料：図 A.8, 図 A.9）

95% family-wise confidence level

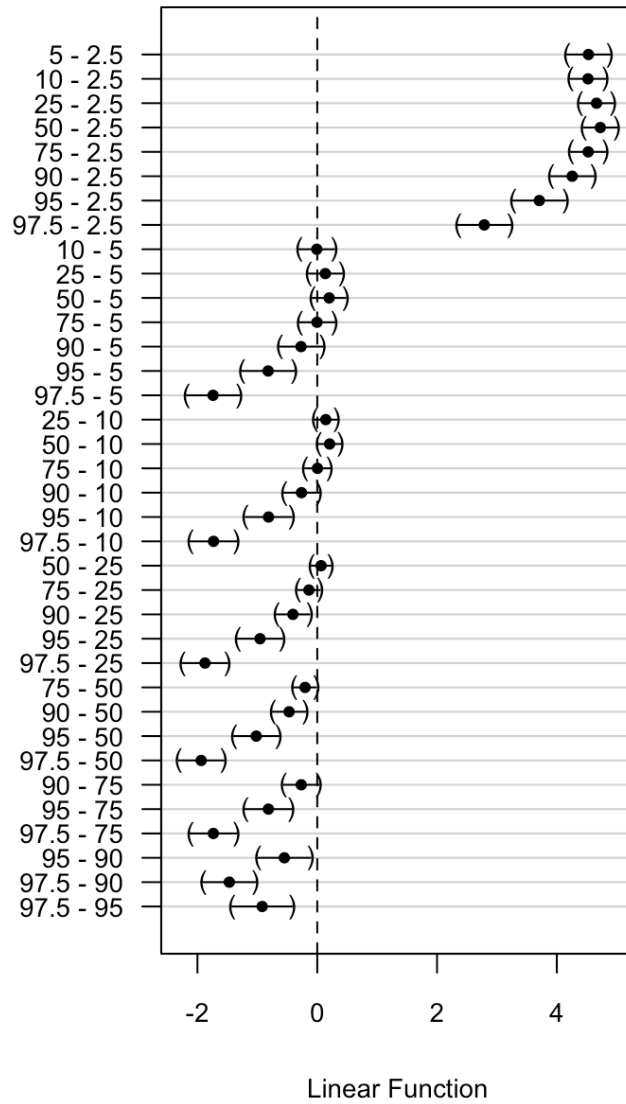


図 3.14: 新陳代謝度分位水準での存続年数の 95%信頼区間

3.3.4 非負値テンソル因子分解による取引変動のパターン分析

前節の新陳代謝度と存続年数に関する有意差検定では、新陳代謝度が高すぎても低すぎても存続年数が短くなるという仮説が支持される結果となった。しかし企業間取引に基づく新陳代謝の概念図 3.1 で示したように、新陳代謝度が増加するパターンは3つあり、これらは提案指標である新陳代謝度単独では区別できない。そこで本節では、これらを区別する企業の取引変動に関する変数を加え、変数間にどのような相関関係があるかを分析する。すなわち、新たに取引数の増加率（企業の成長率）、取引数、消失取引数を加え、新陳代謝度と存続年数と合わせた5つの変数を分析対象とする。分析手法には、高次元データの因子分解技術である非負値テンソル因子分解（NTF）を用いる。

(1) 分析実験データ

本節では TSR の企業間取引データを用いる。取引の変動と企業パフォーマンスの複合的な関係を理解するため、変数として、存続年数、取引増加率、取引数、新陳代謝度、消失取引数を使用した。最初の2属性はカテゴリー値であり、後の4属性は連続値である。存続属性は10年間存続したかどうかのバイナリ値であり、2007年に存在していた企業が2016年にも存在していた場合には1、2007年に存在していた企業が2016年に存在していない場合は0としている。成長属性は企業の取引数変動の傾きが正の場合は1、負の場合は-1、変わらない場合は0としている。連続値についてはカテゴリー値に変換するため、データをソートし、表 3.12 に示すパーセンタイルで区切って集計を行なった。

表 3.12: 変数のカテゴリーと分位点における閾値

Features	Categories / Percentiles / Thresholds
Survival	Not survived, Survived
Growth	Negative, Unchanged, Positive
Transactions p	5, 10, 25, 75, 90, 95, 100
Transactions t	1, 1.75, 3, 10, 15.1, 18.67, 52.4
Metabolism p	10, 25, 75, 90, 95, 100
Metabolism t	0, 0.03, 0.18, 0.28, 0.36, 1
Released p	25, 75, 90, 95, 97.5, 100
Released t	0, 1, 4, 10, 20, 2560

p: percentiles, t: thresholds

それぞれのパーセンタイルに含まれるデータ量は、パーセンタイルに含まれる全データ値が同じでない限り同数である。テンソル X に含まれる要素の値は、各属性のクラスないしパーセンタイル領域に含まれる企業数を正規化したものである。たとえば3次元テンソルで考えた場合、行列 X の要素 x_{ijk} は、生存属性の i 番目のカテゴリー、成長属性の j 番目のカテゴリーおよび取引数属性の k 番目のパーセンタイルに属する企業の数と正規化した値となる。

本節では最適な基底数を求めるため、基底数の数を 2~12 まで振って NTF を実行し、近似誤差の飽和度と減少度を分析した。その結果、繰り返し数 100 回でどの基底数においても近似誤差が飽和し（図 3.15）、基底数 7 以上で近似誤差がほぼ最小値をとることがわかった（図 3.16）。以降の実験では繰り返し数を 100 回、基底数を 7 に定めて NTF を実行する。

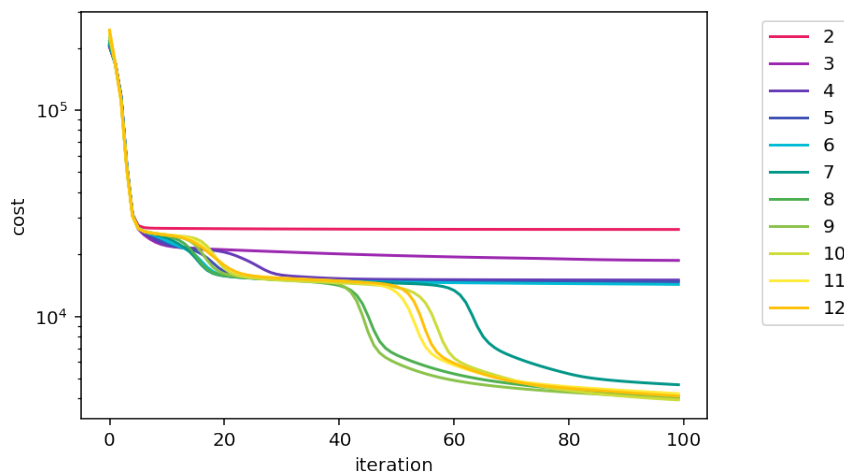


図 3.15: NTF の近似誤差と繰り返し回数

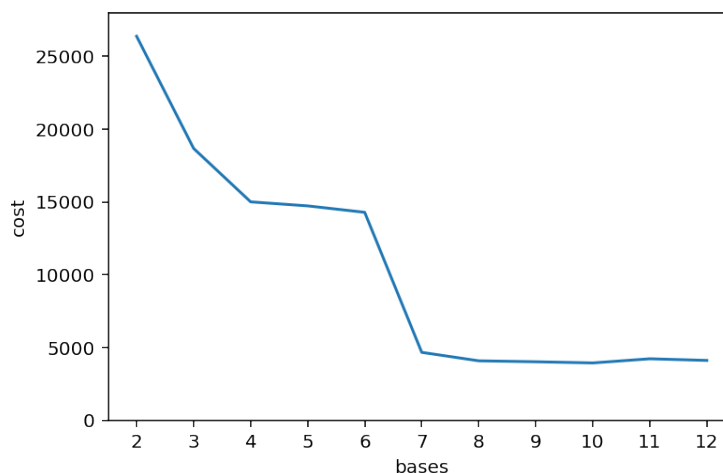


図 3.16: NTF の近似誤差と基底数

(2) NTF による企業属性因子分解結果

図 3.17 は 7 つの基底に因子分解を行なった結果である。X 軸は 5 つの属性を表しており、Y 軸は 7 つの基底それぞれにおける企業数の分布を示している、各基底に属するグループを基底名の順に上からグループ 1,2,3,4,5,6,7 とする。10 年間継続しやすい傾向が見られたの

はグループ6,7であった。最も10年存続率が高いグループ6は取引数がやや増加傾向にあり、取引数が中規模で、新陳代謝度は中程度、消失取引数は少なかった。グループ7は取引数が不変で少なく、新陳代謝度が非常に小さく、消失取引数は少ない傾向があった。これに対し、10年間存続しない傾向が強かったのはグループ1,2,3,4であるが、取引関係の特徴は互いに異なっていた。グループ1は取引数が変動せず、少数の固定的取引を持つ企業で構成され、グループ2に属する企業は取引数の増減があり、比較的少数の、新陳代謝度が高い取引関係を持っていた。グループ3,4の企業は取引数の増減があり、多数の取引先を持ち、新陳代謝度は高く、消失取引数はかなり高い傾向があった。

存続年数に関して例外的な傾向を示したのは、グループ5である。グループ5は減少傾向にある中程度の取引数と中程度の新陳代謝度を持ち、消失取引数が高い傾向があるにもかかわらず、一定程度10年間存続する企業が含まれていた。これらの中規模企業の中には、効率的な取引先の縮減を行いながら、業務改革を行う中で存続し続けた企業群が含まれている可能性がある。

NTFの分析結果を読み解くと、企業間の取引変動と企業の存続年数に関して上記のようなくつかの傾向が浮かび上がってくる。これらに基づくより詳細な分析は今後の研究課題としたい。

3.4 考察と結論

第3章では大規模な企業間取引データを用いて、企業の新陳代謝を定量的に測定することに初めて成功した。2010年から2014年まで5年間の東北・中部・九州地方における企業間取引データを用いた分析により、大半の企業が取引先企業を毎年1~2割程度変更していること、新陳代謝度が低く取引が固定的な企業、あるいは逆に新陳代謝度が極端に高い企業は存続年数が短く、市場から撤退する傾向があること、宮城県や福島県など大企業の多い地域に新陳代謝度の高い企業が多いことが判明した。東日本大震災後のサプライチェーンの崩壊は企業間取引に大きな影響を及ぼしたことが知られているが[170]、被災地である東北地方以外の中部・九州地方でも同様の相関関係が見られたことは、取引先交代による新陳代謝が危機対応時の特異な現象ではなく、平素の企業活動においても不可欠な生存維持機構である可能性を示している。

第3章後半では、2007年から2016年までの10年間の東北地方の企業間取引データを用いて、新陳代謝度と存続年数の間に統計的な有意差があることを確認した。さらに、本章で定義した新陳代謝度では捉えきれない取引変動に関する情報を反映する変数を3つ加え、高次元データの因子分解技術を用いてどのような取引関係の特徴を持つ企業が長く存続するかについて考察した。

東日本大震災前後を含む5年間の東北地方という厳しいビジネス環境において、サプライチェーンの崩壊および再建は企業の存続に大きな影響を及ぼしたことが知られている。産業別・地域別に比較したところ、大規模な復興需要が企業の新陳代謝に影響を与えていることが示唆された。企業間取引の新陳代謝に関する本章の客観的な解析結果は、ミクロな観点で

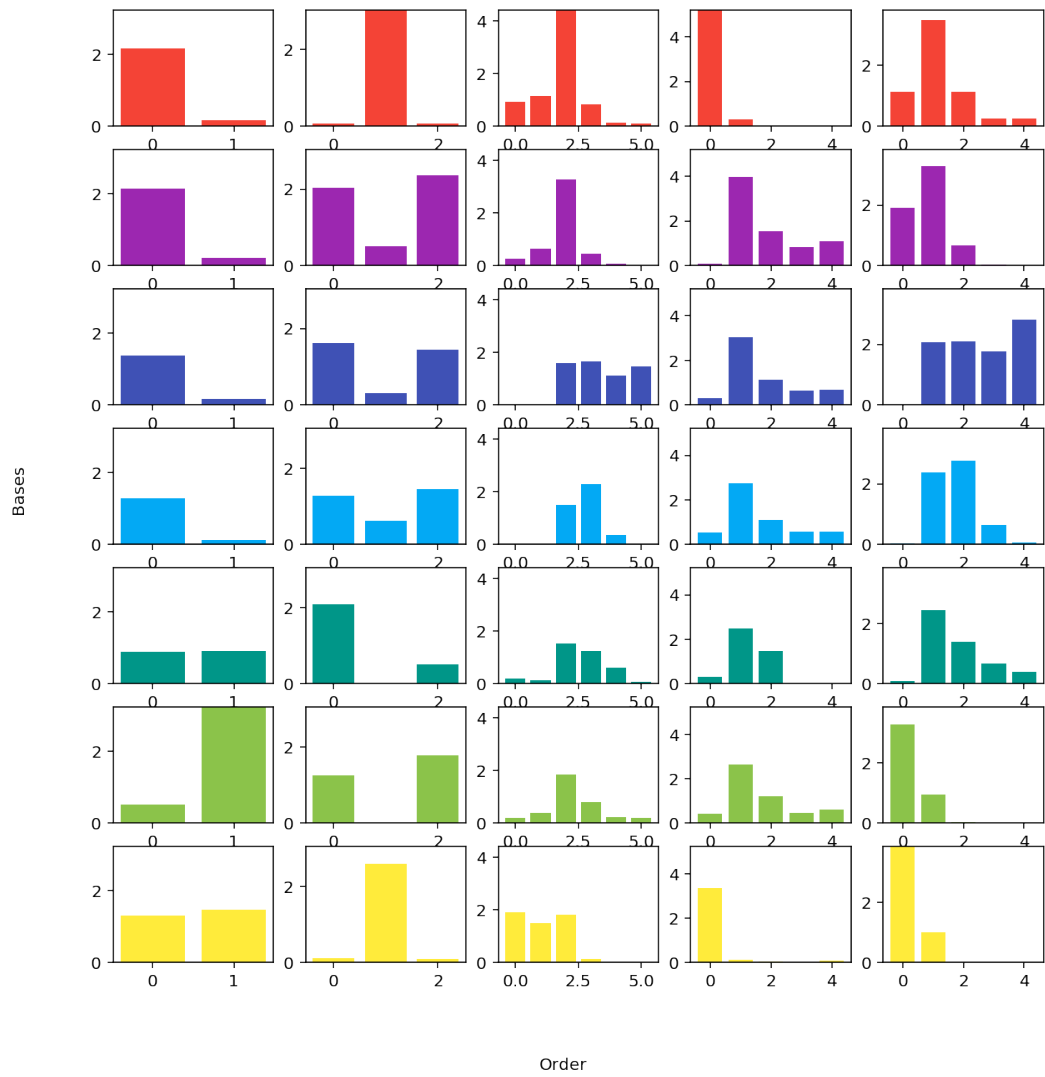


図 3.17: NTF 因子分解結果A:生存 B:成長 C:取引数 D:新陳代謝度 E:消失数

サプライチェーンの変動を捉え直し、環境変化への適応力が企業の新陳代謝度に反映されることを示唆したと言える。

第4章 提案指標2：コミュニティ異質性を反映したノード指標PW

第2章では、オープン化が進み、多様な知見や情報の融合が鍵を握るようになってきたイノベーション活動において、コミュニティ間をつなぐ媒介的な機能を果たす企業の役割が注目されていることを述べた。しかし、製品や人々の嗜好、ニーズの多様化が進んだ社会状況を考慮すると、ただ産業集積間をつなぐだけのコネクターでは、現代的イノベーションの担い手としては十分ではない。今後は関連性の遠いコミュニティを媒介し、新たな組み合わせを生み出すポテンシャルを持つ企業に目を向けることが成功の鍵となる可能性がある。

ネットワーク科学において、遠距離からの知識をもたらす希少なリンクは、スモールワールドにおけるリンクのつなぎ替えで生じるショートカット [171]、異なるコミュニティ間を結ぶ弱い紐帯 [44]、構造的な穴に架かるブリッジ [46] といった概念に相当する。本章では、異質なコミュニティをつなぐ結節点（ノード）を検出する新しいネットワーク指標（PW）を提案する。従来ネットワーク分析におけるノードの重要性は、主にそのノードによって得られる情報量や、隣接関係に基づく中心性の観点から評価されてきた。情報量に基づくノード評価には、ネットワーク内におけるノードを介した情報伝播速度や最短経路数が、隣接関係に基づくノード評価には、ローカルなノード周辺情報が利用されてきたのに対し、ネットワークの重要な特徴であるコミュニティ構造はあまり考慮されてきていない。そこで本章では、コミュニティ間の差異に基づいてリンクの希少性を定量化し、ノード評価に活用することを提案する。また、既存指標では重要なリンクを多数持つノードは評価できても、希少なリンクを持つノードを検出することはできないことを示し、複数コミュニティ間の関連性を考慮する提案指標が、ノードの希少性評価に成功していることを示す [172, 173, 53]。

4.1 比較指標の定義と本章の位置付け

本節では比較指標であるZ値、P値、および提案指標の設計に用いたCRJCについて、それぞれの定義と本章の着眼点からの解釈を述べ、本章の位置づけについて述べる。本章ではノードのコミュニティ介在性について、Guimeraらの提案したP値（participation coefficient P） [47] をベースに提案指標を設計した。P値はノードが持つリンクのコミュニティ網羅性を表す指標であり、コミュニティ間の関係の多様性を反映することはできない。提案指標PWでは、コミュニティ間の関連度でP値を重み付けすることにより、この課題を解決している。

4.1.1 コミュニティ内外のリンクとノードの役割

第2章で説明した媒介中心性, Katz 中心性, Burt の拘束度は, ネットワーク内にコミュニティ構造を想定せず, 着目するノード周囲のノードやリンク構造を用いて評価を行うノード指標である. これに対し, 本節で説明する Z 値および P 値は, コミュニティ内のリンクとコミュニティ外のリンクを区分し, その分布パターンをノードの役割分類に活用した指標 [47] と言える. Z 値はコミュニティ内次数の高さ (ハブ度) を表し, 式 (4.1) で定義され, P 値はコミュニティ間リンク度 (コネクター度) を表し, 式 (4.2) で定義される.

$$Z_i = \frac{K_{S_i} - \bar{K}_{S_i}}{\sigma_{K_{S_i}}}, \quad (4.1)$$

$$P_i = 1 - \sum_{s=1}^{N_m} \left(\frac{K_{S_i}}{K_i} \right)^2. \quad (4.2)$$

K_{S_i} はノード i のコミュニティ S_i 内次数, \bar{K}_{S_i} はコミュニティ S_i の平均次数, σ はコミュニティ S_i 内次数の標準偏差である. N_m はネットワーク内のコミュニティ数であり, K_i はノード i の総次数である. Z 値は標準化されたノード i のコミュニティ内リンク度となる. P 値は二乗和によりリンクの分散を反映する. コミュニティ外リンクを持つほど大きくなり, リンクが全コミュニティに均一に分散していれば 1 に近づき, 所属コミュニティ内にしかリンクがなければ 0 となる. P 値はノードの持つリンクのコミュニティ網羅性を評価する指標であり, コミュニティ間の差異を区別しているわけではない.

Guimera らは大腸菌の代謝ネットワークにおいて, 80% のノードはコミュニティ (彼らの用語ではモジュール) 内にしか反応経路を持たないこと, 異種間で共通に見られる進化的保存性の高いノードは, コミュニティ内次数の高いハブよりも, 異なるコミュニティをつなぐコネクターに多いことを見出した [47]. ここで, P 値の高いノードであるコネクターの特徴は, 複数のコミュニティで共通に用いられるというコミュニティ網羅性にあり, Granovetter や Burt が言うところの異質な情報をもたらすという点にはない. Guimera らは後に隣接ノードの役割にも範囲を広げ, どのような役割を持つノードがどのような役割を持つノードとリンクするかという role-to-role に着目した分類も行っている [157]. しかし role-to-role 分析は, 2パス先までのノード周囲情報を参照した分析であり, ここでもコミュニティ間の差異はノード評価に反映されていない.

4.1.2 コミュニティの構造情報を用いたリンク予測

従来, リンク予測においては共通隣接ノード数やノード類似度といったローカルな構造情報に基づく研究が盛んであったが, 近年ではコミュニティの構造情報がミッシング・リンクの予測に大きく寄与することが明らかになっている [174][175][176][177]. Yan は同一コミュニティ内にあるノードペアは, 違うコミュニティにあるノードペアよりも類似度が高く, つながる確率が高いことを活用し, リンク予測の精度向上に成功した [174]. しかし, ノード

間の直接リンクのみを考慮するこの手法によると、リンクが1つも存在しない場合、両コミュニティをつなぐリンクの存在確率は0になる。Canninskiらは、従来のノードベースの視点にリンクベースの視点を融合させたリンク予測手法であるCAR指標を提案した[175]。この指標は以下の式(4.3)で表され、共通隣接ノード (common neighbor: CN) がローカルコミュニティ内で強いつながり (local-community-links: LCL) を持っている場合に、ノード間のリンク確率が高くなる。CAR指標はLCLをCNの重みとして用いることで、CNや優先的選択 (preferential attachment)[178]、Adamic-Adar[179]といったクラシカルなノードベースのリンク予測より、識別力を高めることに成功している。

$$\begin{aligned} CAR(x, y) &= CN(x, y) * LCL(x, y) \\ &= |\Gamma(x) \cap \Gamma(y)| * \sum_{w \in \Gamma(x) \cap \Gamma(y)} \frac{|\gamma(w)|}{2} \end{aligned} \quad (4.3)$$

x と y はネットワーク中のノード、 w は x と y の共通隣接ノードである。 $\Gamma(x)$ は x の隣接ノードの集合であり、 $|\Gamma(x)|$ は x の次数に相当する。 $|\gamma(w)|$ は、 w の隣接ノードの部分集合である、ローカルコミュニティの次数である。

コミュニティ内の直接リンクのみを考慮した上記手法に対し、コミュニティ間の潜在的関連性を活用したリンク予測手法がDingらによるCom_Stである[177]。Dingらが用いた4つのコミュニティ関連度はいずれもローカルなノード類似度から求めたリンク予測結果よりも高い精度を達成したが、そのなかでも安定的により高いAUCを獲得した尺度は、以下の式(4.4)で定義される、コミュニティ間の共通隣接ノードから算出したジャガード係数 (Community Relevance Jaccard Coefficient: CRJC) である。

$$CRJC(c_i, c_j) = \frac{|\Gamma(c_i) \cup V(c_i) \cap (\Gamma(c_j) \cup V(c_j))|}{|(\Gamma(c_i) \cup V(c_i)) \cup (\Gamma(c_j) \cup V(c_j))|} \quad (4.4)$$

c_i, c_j は $C = \{c_1, c_2, \dots, c_m\}$ で定義されるネットワーク内のコミュニティのいずれかの組み合わせを示しており、 $V(c_i) = \{v | v \in V, c_i \in C(v)\}$ はコミュニティ c_i 内の全てのノードである。 V はネットワークの全てのノード、 $C(v)$ は v を含む全てのコミュニティである。 $\Gamma(c_i)$ は c_i の隣接ノードであり、 c_i 内の全ノードの隣接ノードのユニーク集合となっている。CRJCの分子はコミュニティ単位での共通隣接ノード数であり、分母は2つのコミュニティを構成するノードおよびその隣接ノードの和集合である。

コミュニティ関連度CRJCの概念図を図4.1に示す。赤のノードはコミュニティ c_i, c_j 間の直接リンクに関わるが、青のノードはコミュニティレベルの共通隣接ノードであり、コミュニティ c_i, c_j 間の間接的なリンクに寄与している。

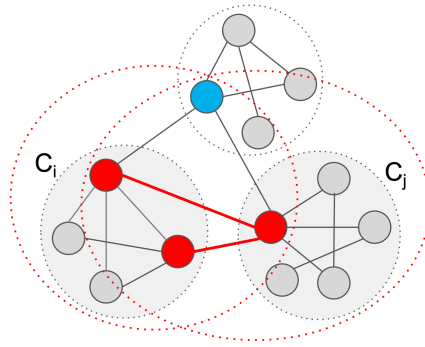


図 4.1: コミュニティ関連度の概念図

4.1.3 本章の位置づけ

本章の仮説は、ノードの重要性は、ノードがもたらすリンクの不均質性から推定されるというものである。多数のリンクを持つノードであるハブが重要であることは既に分かっているが [180]、ノードを重要にするリンクの構成や値についてのエビデンスは比較的少ない。従来のネットワーク指標の多くは、ノードを経由した最短パスの数をカウントする媒介中心性 [150]、一対のノード間のパスの総数を考慮する Katz 中心性 [154]、被リンク (Backlink) 数とその重要度を考慮する Page Rank [155] など、隣接するノードのリンク密度に影響される傾向がある。これらのノード指標は、多数の重要なリンクを持ち、ネットワークへの影響力が大きいノードを抽出するのに有効であるが [180]、少数の重要なリンクを持つ、稀少なノードを抽出することはできない。

本章では、ネットワークの構造的空隙を埋める稀少なノードを検出する新たな指標 PW を提案する。この際、比較指標では活用されてこなかったコミュニティの構造情報を利用する。近年試みられているコミュニティ関連度を用いたリンク予測は、関連性の高いコミュニティであればリンクする確率が高いと考える点で、ノード類似度に基づいて同類を推薦する従来からの予測手法と設計思想が同じである。しかしこうした手法では、スモールワールドや弱い紐帯、構造的空隙といった概念で重視されてきた異質な情報をもたらすノードは検出されない。そこで本章では異質性をもたらすコミュニティ構造に着目し、ノード評価に活用することを提案する。

4.2 提案指標の定義と分析手法

4.2.1 コミュニティ異質性を反映したノード指標

本節では、ノードが持つリンクのコミュニティ網羅性を用いてネットワーク内のノードの役割を分析する指標である P 値 [47] に対する重みとして、コミュニティ関連度 CRJC [177]

を適用し、新たなノード指標 PW を提案する．重み付けのスキームとして、ノードが持つ全てのコミュニティ間リンクの平均 CRJC の負の対数を利用する．これは、ある 2 つのコミュニティが共通隣接ノードを持つ確率によって、コミュニティを結びつけるリンクの希少性を評価するということである．コミュニティ関連度の負の対数を P 値と掛け合わせることで、コミュニティ異質性を反映させた指標 PW を設計した．定義は式 (4.5) である．

$$PW_i = -P_i \log \sum_{j \in \Gamma(i)^{IC}, j \neq i} \frac{CRJC(c_i, c_j)}{L + \delta} \quad (4.5)$$

P_i は式 (4.2) で定義された P 値であり、 P_i の重みは、式 (4.4) で示したコミュニティ関連度の平均の対数を負にした値である． $\Gamma(i)^{IC}$ はノード i の隣接ノードのうち、ノード i の所属クラスター以外に所属している隣接ノードを、 L は $\Gamma(i)^{IC}$ の数を表す．コミュニティ網羅性が高く、関連度の低いコミュニティをつないでいるノードであるほど、 PW_i は大きくなる．

4.2.2 隣接行列の更新による計算行程

P 値の重みは、エッジリストから取得した隣接行列のコミュニティ外リンクに相当する要素を、隣接ノードのコミュニティID、コミュニティ関連度と更新していくことで得られる．図 4.2 に更新行程のアウトラインを示す．なお、計算過程で得られた行列は全てグラフのノード順に対応させている．

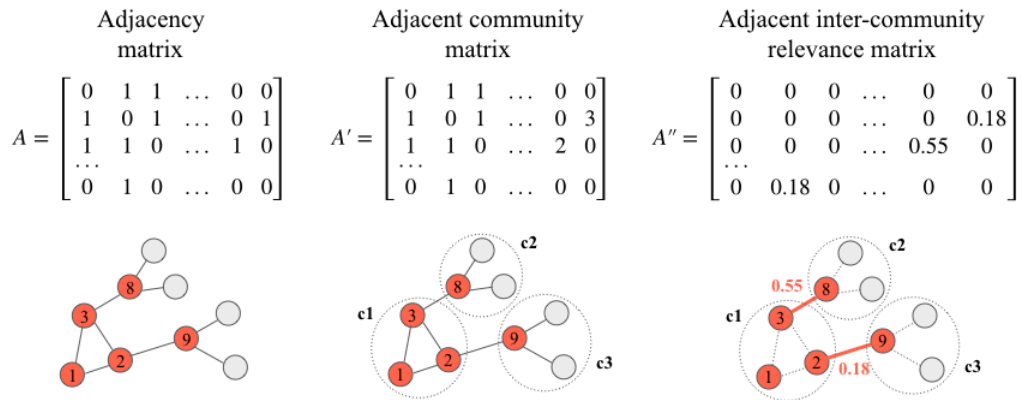


図 4.2: 隣接行列の更新によるコミュニティ関連度取得

(1) 隣接コミュニティ行列

エッジリストから隣接行列を作成する．モジュラリティベースのクラスタリング手法として広く用いられている Louvain 法によりコミュニティ分割を行い、グラフ上の全ノードの所属コミュニティIDを取得する．隣接行列と対角成分に所属コミュニティIDが入った対角行

列の内積を取ることで、リンクの存在する要素位置に隣接ノードのコミュニティIDが入った隣接コミュニティ行列を作成する。

(2) 非所属隣接コミュニティ行列

各コミュニティのノードIDを取得する。隣接コミュニティ行列のうち、各ノードの所属コミュニティIDは0に置換し、所属コミュニティの外にある隣接ノードのコミュニティIDのみを残す。その際、隣接コミュニティ行列の要素を、各コミュニティのノードIDで指定する作業を繰り返すことで、全てのコミュニティについて同様に、所属コミュニティの外にある隣接ノード（非所属隣接ノード）のコミュニティIDのみを抽出する。得られたコミュニティID行列を連結し、非所属隣接コミュニティ行列とする。

(3) 非所属隣接コミュニティ関連度行列

グラフ上の全てのコミュニティペアについて、計算式(4.4)に基づいてコミュニティ関連度を取得する。非所属隣接コミュニティ行列のコミュニティIDをコミュニティ関連度に置き換えた、非所属隣接コミュニティ関連度行列を作成する。その際、各ノードについて、所属コミュニティと非所属隣接ノードの所属コミュニティとの関連度を求める作業を、ノードの所属コミュニティ毎に繰り返すことで、全てのコミュニティについて同様に、所属コミュニティの外にある隣接ノードのコミュニティ関連度を抽出する。得られたコミュニティ関連度行列を連結し、非所属隣接コミュニティ関連度行列とする。

(4) 平均非所属隣接コミュニティ非関連度

各ノードのコミュニティ外次数を求める。非所属隣接コミュニティ関連度行列から、ノード毎に非所属隣接コミュニティ関連度の合計を求め、コミュニティ外次数で割ることで、コミュニティ関連度の平均値を取得する。この際、ノードのコミュニティ外次数が0である場合はinf回避のため、分母に $\delta = 0.000001$ を加える（分子のコミュニティ関連度が0なので結果的に0になる）。平均コミュニティ関連度の対数を負にした値を非関連度とする。平均コミュニティ関連度が0の場合は、0を1に置き換えた値を非関連度とする。

(5) 重み付きP値 (PW)

各ノードの次数を取得する。隣接コミュニティ行列を用いて、計算式(4.2)に基づき、ノード毎に各コミュニティに属する隣接ノード数をノードの次数で割った値の二乗の総和を求めることでP値を取得する。P値に平均非所属隣接コミュニティ非関連度を掛けることで、ノード毎の重み付きP値を取得する。

P値の重みに用いた平均非所属隣接コミュニティ非関連度の計算手順を以下にまとめる。

1. 隣接行列から隣接コミュニティ行列を作成する
2. コミュニティ外隣接ノードのコミュニティID を抽出する
3. 全てのコミュニティペアについてコミュニティ関連度を計算する
4. コミュニティ外隣接ノードのコミュニティID をコミュニティ関連度で置換する
5. 各ノードのコミュニティ外次数を計算する
6. 各ノードについてコミュニティ外隣接ノードのコミュニティ関連度を合計し、コミュニティ外次数で割ることで平均コミュニティ関連度を計算する
7. 平均コミュニティ関連度の対数を負にした値を非関連度とする

4.2.3 P 値と PW 値のノード識別力の比較

図 4.3 では SNAP[54] が提供する Facebook の友達関係データを用いて、P 値と PW 値のコミュニティ介在性に関するノード識別力を比較した。このデータの中には 10 人の匿名ユーザーのエゴネットワークと友人リストが含まれている [181]。中央の棒グラフは G0-G9 までの 10 人の Facebook ユーザー毎のエゴネットワークと 2 種類の統合ネットワーク（全ユーザーの統合ネットワークおよびユーザー G0, G2, G6 の統合ネットワーク）において、P 値（灰色）と PW 値（緑色）がとりうる値の範囲を示している。P 値と PW 値の平均値はそれぞれ灰色と緑色の円で示した。P 値、PW 値の平均値は比較的近傍にあるが、値のとりうる範囲は上位層において PW 値の方が広いことがわかる。なお、P 値・PW 値の詳細なデータ分布は付録 A.15 のバイオリンプロットにも示しており、ここでも上位層において PW の表現力が高い傾向が確認されている。

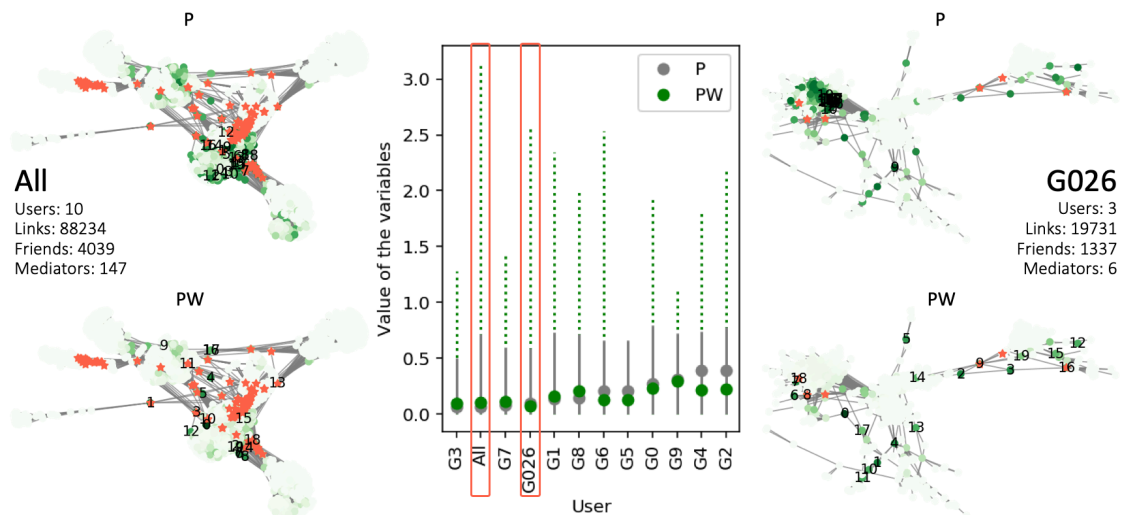


図 4.3: P 値と PW 値のノード識別力の比較

両サイドのネットワーク図は2種類の統合ネットワーク上でのP値およびPW値の比較である。図中には各ネットワークのユーザー数、リンク数（友達関係の数）、ノード数（友人数）、介在ノード数（ユーザーペア毎の共通の友人数の合計）が示されている。上位20番目までのノードには順位を表す数値を付けており、ノードの緑色の濃さは各ネットワークでのノード値に対応している。全てのネットワークは互いに関連度が異なるクラスターで構成されており、橙色の星形ノードで示されたエゴユーザー間の共通の友人は、コミュニティの間に位置している。ここで、P値の高いノードがコミュニティが密集している領域に凝集しているのに対し、PW値の高いノードは離れたコミュニティの間に位置する傾向があり、稀少な共通の友人を見つける上でP値よりよいパフォーマンスを示している。なお、各ネットワークにおけるコミュニティの位置関係は補足資料として図A.13、図A.14に、上位20ノードのP値およびPW値は表A.8に示した。

4.3 検証実験

4.3.1 データセット

本節では指標の有効性を検証するため、ベンチマークデータとしてRCネットワークとLFRネットワーク、実データとして東北地方の企業間取引ネットワークを使用し、評価実験を行った。以下それぞれのデータセットについて説明する。

(1) RC ネットワーク

Watts が提唱した RC (Relaxed caveman) ネットワークは小さい L と大きい C、大きすぎない平均次数を同時に満たすスモールワールド性を持つ [171]。本節では、コミュニティ関連度の差異が大きいできる限りシンプルなネットワークを生成することを目的とし、パラメータであるコミュニティ数、コミュニティ内ノード数、繋ぎ替え確率（リワイヤリング率）、シードはそれぞれ 3, 8, 0.2, 16 に設定した。リワイヤリング率はコミュニティ間のリンク率を指しており、リワイヤリング率が大きいほど密につながったコミュニティとなる。可視化のために生成したネットワークのノード数は 24、エッジ数は 84、コミュニティ数は 3、平均最短経路長は 2.073、平均クラスター係数は 0.676 であった。攻撃耐性の実験では、コミュニティ数 2、3、12 の RC ネットワークを 3 つ生成し、各コミュニティのノード数は 8、リワイヤリング率は 0.2 とした。

(2) LFR ネットワーク

Lancichinetti らが提唱した LFR (Lancichinetti-Fortunato-Radicchi) ネットワークは現実のネットワークに近い、柔軟できめ細かなパラメータ設定が可能なベンチマークネットワークである [48]。Relaxed caveman ネットワークではコミュニティ内のノード数が全て同じ値に固定されていたのに対し、LFR ネットワークでは、ノード数の異なるコミュニティ

を生成することができ、最大次数や冪指数に基づいて次数分布を指定することもできる。パラメータは下記に設定した。-N 100 -k 8 -maxk 30 -mu 0.1 -minc 5 -maxc 50 これはノード数 100, 平均次数 8, 最大次数 30, リワイヤリング率 0.1, コミュニティ内ノード数 5~50 のネットワークを生成することを示している。次数分布の冪指数はデフォルトで 2 に, コミュニティサイズの冪指数はデフォルトで 1 に設定されている。可視化用に生成したネットワークのエッジ数は 391, コミュニティ数は 5, 平均最短経路長は 2.804, 平均クラスター係数 0.492 であった。攻撃耐性の実験では, ノード数 100, ノード毎の度数 8, リワイヤリング率 0.1~0.6 までの 6 つの値をとる LFR ネットワークを生成した。なお, 0.5 を超えるリワイヤリング率では, ネットワークからコミュニティ構造が失われる。

(3) 東北ネットワーク

東北ネットワークは株式会社東京商工リサーチ (TSR) から提供を受けた, 2016 年の東北地方の企業間取引データを使用した。取引データ総数は 154,918 件であったが, 次数分布の極端な偏りを避けるため, 2000 以上のエッジを持つノードと 10 未満のエッジを持つノードを除外し, ノード数 (企業数) 317, エッジ数 (取引数) 2412 のデータを取得した。東北ネットワークでは, コミュニティ間のリンクが少ない 1 つのコミュニティと, コミュニティ間のリンクが多い 4 つのコミュニティが存在していた。

4.3.2 実験内容と結果

(1) コミュニティ関連度

使用した 3 つのネットワークのコミュニティ関連度 CRJC を表 4.1, 表 4.2, 表 4.3 に示す。表中の Paris は各ネットワークにおけるコミュニティ ID の組み合わせを表しており, Relevance は各コミュニティペアにおける関連度を示している。3 つのネットワークはそれぞれ異なるコミュニティ関連度の分布傾向を示しており, 下記の 3 つのタイプに分けられる。

- 全コミュニティペアでほぼ均等な関連度差
RC ネットワークはコミュニティ 1 と 2 の関連度が最も高く, コミュニティ 1 と 3 の関連度が最も低かった。各コミュニティペアにおける関連度の差はほぼ均等で 0.2 程度であった (表 4.1)。
- 高関連度な 1 対のコミュニティペア以外全て低関連度
LFR ネットワークはコミュニティ 2 と 4 の関連度が最も高く, それ以外のコミュニティペアの関連度は低かった。コミュニティ 1 と 3, およびコミュニティ 3 と 5 は, 共通隣接ノードが存在しないため関連度が 0 であった。(表 4.2)。

- 1つのコミュニティを除き全て高関連度
東北ネットワークではコミュニティ1の関連度が低く、それ以外の全てのコミュニティペアの関連度が高かった。(表 4.3)。

表 4.1: RC ネットワークにおけるコミュニティペアの関連度

Pairs	Relevance
(1, 2)	0.550
(1, 3)	0.182
(2, 3)	0.350

表 4.2: LFR ネットワークにおけるコミュニティペアの関連度

Pairs	Relevance
(1, 2)	0.107
(1, 3)	0.046
(1, 4)	0.0
(1, 5)	0.0
(2, 3)	0.471
(2, 4)	0.225
(2, 5)	0.081
(3, 4)	0.153
(3, 5)	0.032
(4, 5)	0.0

表 4.3: 東北ネットワークにおけるコミュニティペアの関連度

Pairs	Relevance
(1, 2)	0.149
(1, 3)	0.211
(1, 4)	0.104
(1, 5)	0.075
(2, 3)	0.570
(2, 4)	0.555
(2, 5)	0.394
(3, 4)	0.580
(3, 5)	0.449
(4, 5)	0.495

4.3.3 ノード指標別順位とネットワーク可視化

提案指標のパフォーマンスを確認するため、3つのネットワークを可視化し、5つの比較指標と提案指標によるノード値に従って順位付けを行った。可視化した3つのネットワークを図 4.5, 図 4.6, 図 4.7 に示す。ネットワークの可視化にはばねモデル (spring layout) として知られている Fruchterman-Reingold Algorithm [182] を用いた。ばねモデルは互いにリンクしているノードほど近くに描画する性質を持つことから、描画されたコミュニティ間の距離は、共通隣接ノードを元に計算しているコミュニティ関連度を反映していると考えられる。

3つのネットワークをコミュニティ毎に色分けして位置関係を示したのが図 4.4 である。なおコミュニティ分割にはいずれのネットワークにおいてもモジュラリティベースのクラスタリング手法である Louvain 法を適用した。

可視化に用いたノード指標は、Z 値 (Z), Katz 中心性 (Katz), 媒介中心性 (Bet), P 値 (P), PW 値 (PW), そして構造的空隙を表すとされてきた Burt の拘束度の逆数 (iBurt) で

ある。ノードの色は各指標の値を表しており、色が濃いほど値が大きいことを表す。ノード上の番号は各指標におけるノード順位である。LFR および東北ネットワークでは上位 20 位までのノードの周囲を橙色でマークし順位を表示した。3 種類の全てのネットワークにおいて、遠く離れたコミュニティをつなげるノードを識別する力が最も高かったのは PW 値であり、P 値はコミュニティ間をつなぐノードを高く評価したが、コミュニティ間差異は識別していなかった。他の 4 つの指標である Z 値, Katz 中心性, 媒介中心性, iBurt 値は次数の影響を強く受け、コミュニティ間差異は区別していなかった。また、この傾向はネットワークが複雑化し、次数分布が多様になるほど強くなった。以下詳説する。

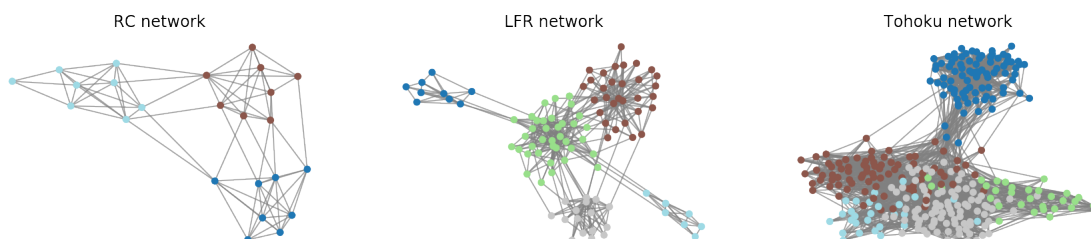


図 4.4: 各ネットワークにおけるコミュニティの位置関係。RC ネットワークは関連度の差がほぼ同じ 3 つのコミュニティ, LFR ネットワークは高関連度な 1 対のコミュニティペア以外全て関連度が低い 5 つのコミュニティ, 東北ネットワークは 1 つのコミュニティを除き全て関連度が高い 5 つコミュニティを持つ。

(1) RC ネットワークの可視化

図 4.5 の RC ネットワーク中で 3 つのコミュニティ全てにリンクを持つノードは 1 つだけであり、そのノードをコネクター、次数の最も大きいノードをハブと呼ぶことにする。本節ではまずはじめに、この 2 つの特徴的なノードを、提案指標を含む 6 つの指標がどのように評価するかを確認した。Z 値以外の全ての指標がコネクターを上位にランクづけているのに対し、Z 値はコネクターを 20 番目にランク付けしている。媒介中心性, P 値, PW 値はコネクターを最上位 (0 番目) にランク付けしており、Katz 中心性と iBurt 値はそれぞれ 4 番目と 1 番目にランク付けしている。これに対し、ハブは全ての指標において上位に位置付けられており、Z, Katz, iBurt 値はハブを 0 番目に、媒介中心性はハブを 1 番目にランクさせている。前節で見たように、これらの指標が次数の影響を大きく受けることは定義上明らかであるが、可視化とランキング結果もこれに一致していることが読み取れる。

次に 6 つのノード指標がコネクターとハブ以外のノード評価において、コミュニティ関連度を適切に反映させているかを確認する。説明のため、RC ネットワークの 3 つのコミュニティの ID を右から順に c1, c2, c3 とする。各コミュニティ間の関連度は表 4.1 に示している。まず、Z 値と Katz 中心性はコミュニティ外リンクを持つノード自体を高く評価できていない (c1 においてその傾向が顕著)。P 値, PW 値, iBurt 値, 媒介中心性はいずれも他のコミュニティとの境目にあるノードの順位が高いが、コミュニティ間の差異を反映してい

るのはPW 値のみである．中央に位置する c2 の上位 4 ノードを見ると，関連度が小さい c2 と c3 をつなぐノードの順位はPW 値で 3 位，4 位であり，関連度が大きい c1 と c2 をつなぐノードの順位はPW 値で 8 位，9 位であることから，PW 値は関連度が小さいコミュニティをつなぐノードをより高く評価していることがわかる．これに対し，P 値と iBurt 値はいずれも，c1 と c2 をつなぐノードと c2 と c3 をつなぐノードの順位差がほとんどないことから，コミュニティ間差異を考慮しない指標であることがわかる．媒介中心性は c2 だけを見るとコミュニティ間差異を区別しているように見えるが，関連度の大きい c1 と c2 をつなぐノードを，関連度の小さい c2 と c3 をつなぐノードより高く評価しており，コミュニティ間差異を適切に評価できていないことがわかる．

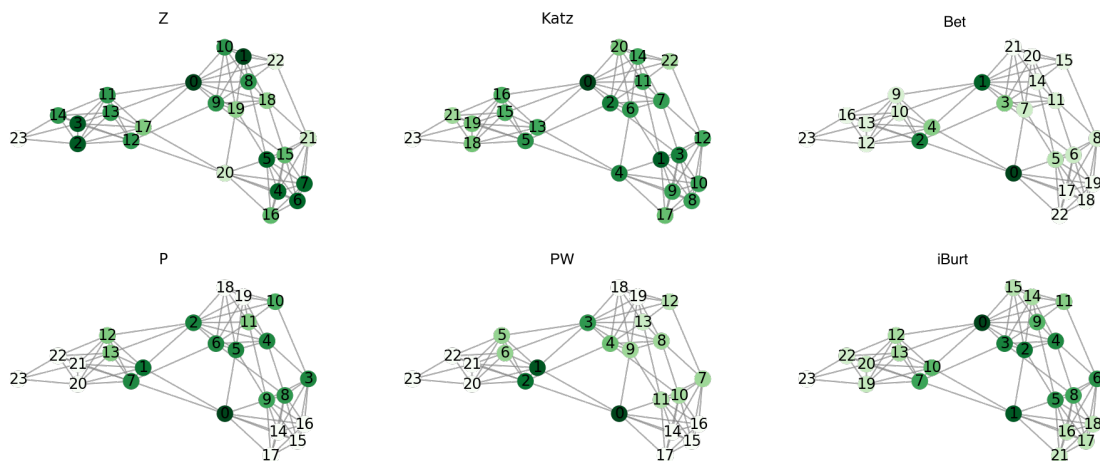


図 4.5: RC ネットワークの指標別ランク. 各ネットワークは 24 ノード，84 エッジ，3 つのコミュニティで生成され，リワイヤリング率は 0.2 である．

(2) LFR ネットワークの可視化

ネットワーク特性に応じた指標の識別力を検証するため，次数分布の異なる 5 つのコミュニティを持つ LFR ネットワークを生成し，ノード指標に基づくランキングと共に可視化した (図 4.6)．各コミュニティ間の関連度は表 4.2 に示した通りである．LFR ネットワーク中の 5 つのコミュニティの ID を右から順に c1, c2, c3, c4, c5 とする．このネットワークでコミュニティ間ノードの順位が高いのは P 値，PW 値のみであった．両指標はノードの持つ全リンク数に対するコミュニティ外リンクの割合を反映するため，次数が小さいノードも高順位になりうる．しかし，P 値では最もコミュニティ関連度の高い c2 と c3 をつなぐノードが高順位であるのに対し，PW 値では，c2 と c3 をつなぐノードの順位は低く，それ以外の関連度の低いコミュニティ間をつなぐノード順位が高い．ここでも遠く離れたコミュニティをつなぐノードを識別するのは PW 値のみであった．

Z 値，Katz 中心性，媒介中心性，iBurt 値はいずれもコミュニティ内部にあるノードの順位が高くなっており，ネットワークの構造的空隙を識別できていないが，興味深い点として，

各指標ごとの評価に特徴的な違いが見られた。以下、定義に基づく解釈と共に説明する。まず、Z 値は全てのコミュニティ内に指標値の高いノードが存在するのに対し、Katz 中心性では最も大きなコミュニティ内に高評価ノードが集中している。前節で説明した通り、Z 値がコミュニティ内ハブを識別する指標であり、Katz 中心性が指標値が最も高いノードの周囲にあるノードを高く評価する指標であることを反映していると言える。媒介中心性は対象ノードを経由する経路数を考慮しているため、他のコミュニティを媒介しているノードが多い c2 に高評価ノードが集中している。iBurt 値は最も大きな 2 つのコミュニティ内に高評価ノードが集中している。iBurt 値は 2 パス先までのノード周囲情報を反映し、3 パス以上先のノード情報は考慮しないため、大半のノードが 3 パス以上のステップを踏まないに到達できない c1 や c5 に存在するノードは評価できないことがわかる。

Z 値、Katz 中心性、媒介中心性、iBurt 値はいずれも次数の影響を受ける指標であり、P 値や PW 値とは違って次数が小さいノードは高順位になりにくい。ノード数とエッジ数を増やし、次数分布をばらつかせてネットワークを複雑化させると、これらの指標では構造的な空隙を捉えることができないことが視覚的にも明らかとなった。

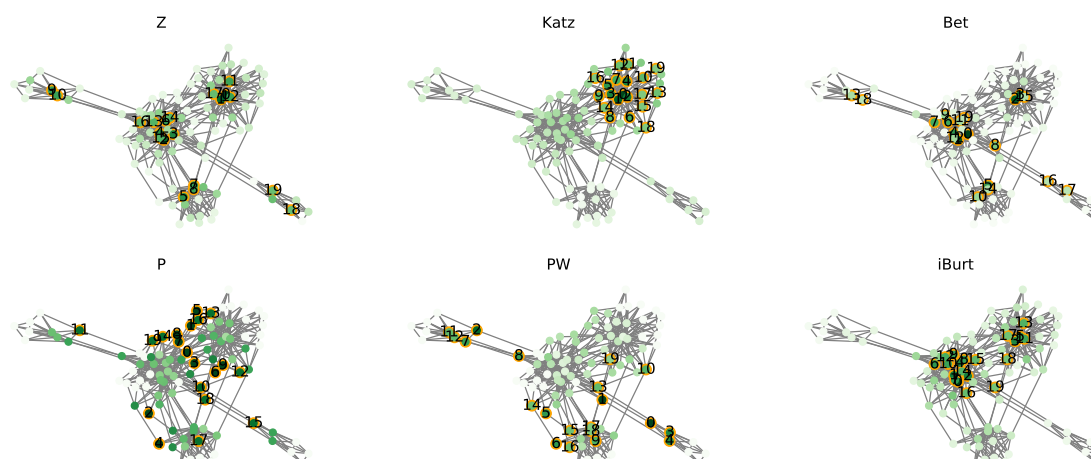


図 4.6: LFR ネットワークの指標別ランク. 各ネットワークは 100 ノード, 391 エッジ, 5 つのコミュニティで生成され, リワイヤリング率は 0.1 である.

(3) 東北ネットワークの可視化

P 値の限界は東北ネットワークにおいてより顕著に現れ、関連度の低い異質なコミュニティをつなぐノードの検出には PW 値が最も優れていることが明らかになった (図 4.7)。東北ネットワーク中の 5 つのコミュニティの ID を上および右から順に c1, c2, c3, c4, c5 とする。コミュニティ c1 のみが他の 4 つコミュニティと関連度が低く、4 つのコミュニティ c2, c3, c4, c5 は互いに関連度が高かった (表 4.3)。このネットワークでは、P 値の高順位ノードは関連度の高い c2, c3, c4, c5 からなる下部のコミュニティ集合内部に集中している一方、PW 値の高順位ノードは関連度の低い c1 と下部の 4 つのコミュニティをつなぐ位置に集中し

ていた。これに対し、Z 値、iBurt 値における高順位ノードは、c1 内部もしくは関連度の高い4つのコミュニティ群の中に集中する傾向、Katz 中心性は c2 に集中する傾向があり、構造的空隙を識別できていなかった。媒介中心性は関連度の低いコミュニティ間をつなぐノードを識別していたが、c1 内部のノードも高く評価しており、ここでも異質なコミュニティをつなぐノードの識別に最も優れていたのは PW 値であった。

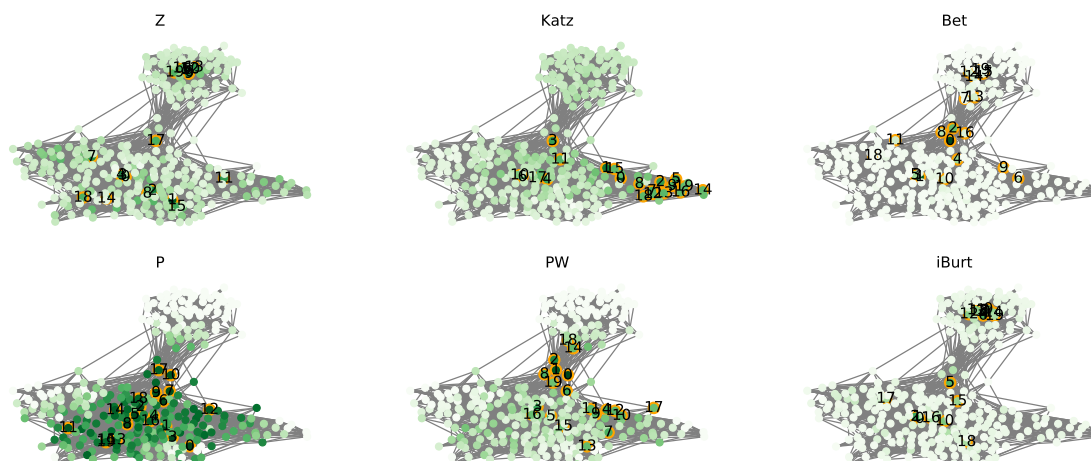


図 4.7: 東北ネットワークの指標別ランク。各ネットワークは 217 ノード、2696 エッジ、および5つのコミュニティで構成される。上部に存在する1つのコミュニティは下部に存在する4つのコミュニティに比べて、コミュニティ間結合が少ない。

4.3.4 ネットワークへの攻撃耐性とノードの順位相関

弱い紐帯や構造的空隙という概念で提示されたコミュニティ間リンクの重要性はそのリンクがもたらす情報の希少性にある。本章では、情報の希少性をコミュニティ間の共通隣接ノードの少なさとして捉え直し、ノード評価に活用した。共通隣接ノードが少なく、関連度の低いコミュニティ同士をつなぐリンクを持つノードを削除した場合、そのノードを介さずに両コミュニティをつなぐリンクの少なさから、ネットワークの平均最短経路長は長くなることが予想される。

次数分布が不均質 (inhomogeneous) なネットワークでは、ネットワークの連結性 (connectivity) に重要な役割を果たすノードを削除すると、ネットワーク耐性が弱まることが知られている [183]。本節では、ネットワークの攻撃耐性はネットワーク内のコミュニティ構造によっても影響を受けることを論証する。コミュニティを識別する3つの指標である Z 値、P 値、PW 値を用いて、攻撃 (ノード削除) に対するネットワークの耐性評価を行う。その結果、攻撃された際により少ないエッジ損失で、ネットワークの連結性を最も大きく損なうノードを識別する指標は PW 値であった。なお、平均最短経路長は媒介中心性が高いノードを削除しても長くなるが、媒介中心性が反映するのはノードを経由する経路数であり、コミュニティ構造ではないため、今回の比較対象からは除外した。

まず、ネットワークの耐性評価の適用方法について説明する。ネットワーク内のノードを順に1つだけ削除した後の平均最短経路長（式(4.6)）を計算し、平均最短経路長の長い順にノードIDを並び替えて標準順位とする。さらに、Z値、P値、PW値の各指標での評価の高い順にノードを1つずつ削除した際の平均最短経路長を計算し、標準順位での順位との違いをSpearmanの順位相関係数で比較した。

$$a = \sum_{s,t \in V} \frac{d(s,t)}{n(n-1)}, \quad (4.6)$$

V はネットワーク内のノード群、 $d(s,t)$ はノード s とノード t をつなぐ最短経路の長さ、 n はネットワーク内の全ノード数である。

(1) コミュニティ数の異なるRCネットワークでの順位相関

図4.8は、左から順に2個、3個、12個のコミュニティを持つRCネットワークでのノード削除実験の結果である。なお、どのネットワークもクラスター内ノード数は8、エッジの繋ぎ換え確率は0.2とした。図の横軸はZ値、P値、PW値でのノード順位、縦軸は標準順位である。Z値、P値、PW値の順位はそれぞれ水色、黄緑、緑のドットで示している。

コミュニティ数が2の場合は、P値とPW値が同じ結果となり、緑と黄緑のドットが重なっていた（図中、P値とPW値が等しいドットは赤い縁でマークしている）。コミュニティ数が3の場合は、中間の順位においてはP値とPW値で差があったが、下位と上位ではP値とPW値の順位が同じであった。コミュニティ数が12の場合は、下位中位上位のいずれにおいてもP値とPW値の順位に差があり、PW値での順位相関がより高い結果となった。なお、Z値は3つのネットワークのいずれにおいてもランダムな分布を示し、標準順位との相関がなかった。

コミュニティ数が2の場合にPW値とP値の順位が等しくなったのは、コミュニティペアが1つしかなく、全ノードにおいて、PW値がP値に同じコミュニティ関連度を掛けた値となったためである。これに対し、コミュニティ数が3の場合に中間順位でP値とPW値が一致しなかったのは、コミュニティ関連度が3つになり差異が生じたためと考えられる。ただし、下位ではP値とPW値の順位が重なり、相関が見られなかった。これは全てのリンクがコミュニティ内にしか存在しないノードはP値が0であり、その結果PW値も0になって順位に意味がなくなるためである。上位ノードでP値とPW値が同じか近くなったのは、これらのノードが3つのコミュニティ全てにリンクを持ち関連度の差が小さくなったためと考えられる。コミュニティ数が12の場合は、コミュニティ外リンクの割合が相対的に高くなり、P値とPW値の順位は一致しなかった。

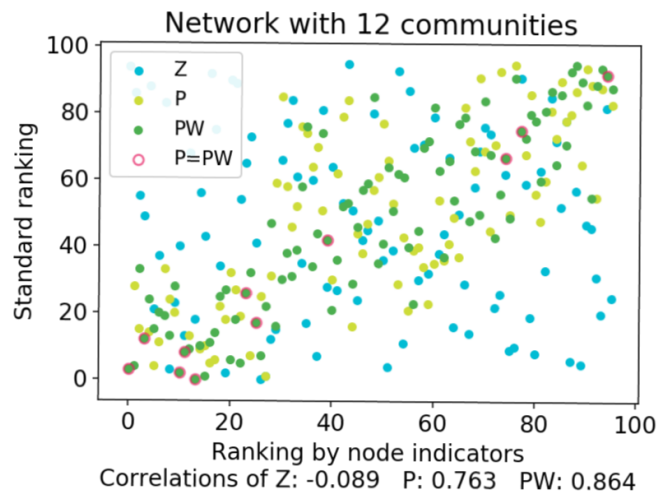
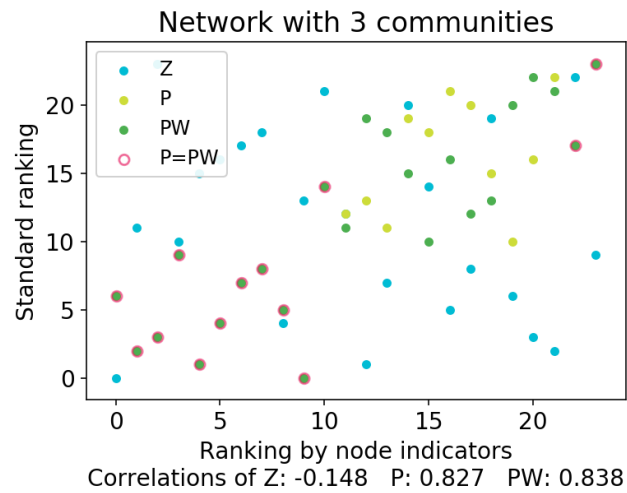
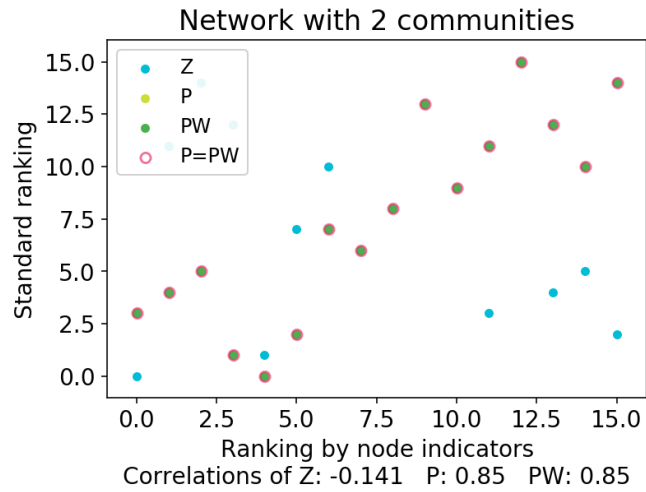


図 4.8: ノード削除後の平均最短経路長による順位相関の指標別比較. 使用した RC ネットワークは左から順に 2, 3, 12 個のコミュニティ, 16, 24, 96 ノード, そして 56, 84, 336 エッジを持つ. 各図の下には標準順位と各指標での順位との Spearman の順位相関係数を示した

次に、2個から17個まで、コミュニティ数の異なる16個のRCネットワークを生成し、同様にして順位相関を比較した(表4.4)。それぞれのネットワークにおいて、各コミュニティを構成するノード数は8であり、リワイヤリング率は0.2とした。表中のCnは各ネットワーク中のコミュニティの数、Z, P, PWで示されたカラムには各指標と標準順位とのSpearmanの順位相関係数である。Z値はいずれのネットワークにおいても順位相関が低かった。P値とPW値はコミュニティ数2の場合には順位相関がどちらも0.85で同じであったが、それ以外のネットワークにおいては全てPW値の順位相関が高かった。

表 4.4: コミュニティ数の異なる RC ネットワークでの順位相関

Cn	Z	P	PW
2	-0.141	0.850	0.850
3	-0.148	0.827	0.838
4	0.048	0.507	0.733
5	-0.174	0.676	0.825
6	-0.194	0.663	0.823
7	-0.208	0.792	0.863
8	0.006	0.727	0.792
9	-0.119	0.737	0.864
10	-0.068	0.785	0.880
11	-0.022	0.711	0.846
12	-0.089	0.763	0.864
13	-0.126	0.75	0.852
14	-0.157	0.746	0.863
15	-0.101	0.706	0.861
16	-0.074	0.779	0.889
17	-0.027	0.775	0.869

(2) コミュニティサイズの異なる RC ネットワークでの順位相関

5個から20個まで、コミュニティを構成するノード数(コミュニティサイズ)の異なる16個のRCネットワークを生成し、同様にして順位相関を比較した(表4.5)。それぞれのネットワークにおいて、コミュニティ数は5であり、リワイヤリング率は0.2とした。Snは各ネットワークにおけるコミュニティサイズ、Z, P, PWで示されたカラムには各指標と標準順位とのSpearmanの順位相関係数である。Z値はいずれのネットワークにおいても順位相関が低かった。P値とPW値では一部コミュニティサイズの大きいネットワークでP値の順位相関がわずかにPW値を上回ったが、総じてPW値の順位相関が高かった。

表 4.5: コミュニティサイズの異なる RC ネットワークでの順位相関

Sn	Z	P	PW
5	-0.597	0.735	0.749
6	0.238	0.757	0.871
7	0.064	0.718	0.871
8	-0.174	0.676	0.833
9	-0.159	0.844	0.888
10	-0.066	0.869	0.893
11	0.071	0.847	0.909
12	-0.182	0.903	0.93
13	-0.284	0.914	0.912
14	-0.09	0.828	0.859
15	0.014	0.861	0.894
16	-0.187	0.886	0.89
17	-0.16	0.857	0.856
18	-0.21	0.842	0.882
19	-0.186	0.888	0.903
20	-0.129	0.881	0.887

(3) リンク率の異なる LFR ネットワークでの順位相関

0.1から0.6まで、リワイヤリング率（コミュニティ間リンク率）の異なる6個のLFRネットワークを生成し、同様に順位相関を比較した（表4.6）。それぞれのネットワークにおいて、コミュニティ数は10、ノード数は100、ノードあたりの平均次数は8である。Muは各ネットワークにおけるリワイヤリング率、Z,P,PWで示されたカラムには各指標と標準順位とのSpearmanの順位相関係数である。これらのネットワークではいずれの指標においてもリワイヤリング率が低くなるほど、標準順位との順位相関が低くなる傾向があった。リワイヤリング率の小さいネットワークでZ値は弱い負の順位相関を示したが、いずれのネットワークにおいてもP値、PW値よりは低い順位相関であった。また、PW値は全てのネットワークにおいて最も順位相関が高かった。コミュニティ間リンク率が上がるほど順位相関が低くなったのは、P値、PW値がともにコミュニティ間リンクを評価する指標だからだと考えられる。コミュニティ間リンク率が低いほどコミュニティ間差異は大きくなるため、PW値はP値より優位であった。コミュニティ内リンクよりもコミュニティ外リンクの数が多くなり、コミュニティ構造が消失するリンク率0.5、0.6では、P値、PW値ともに標準順位との相関が見られなかった。

以上、表4.4、表4.5、表4.6において示された、コミュニティ数Cn、コミュニティサイズSn、コミュニティ間リンク率Muの異なる、合計38種類のネットワークにおけるノード削除実験の結果をまとめる。Z値はいずれのネットワークにおいても順位相関が低かった。全

でのネットワークにおいて PW 値の順位相関は P 値の順位相関より高いかほぼ同じであった。このことは、削除されたときにネットワークの平均最短経路長を長くするノードを識別する力が、P 値より PW 値の方が強いことを指している。これにより、PW 値は他に代替の効かない希少なノード、すなわちそのノードを介さないと到達できない経路を、P 値よりも正確に評価していることが示された。

表 4.6: リンク率の異なる LFR ネットワークでの順位相関

Mu	Z	P	PW
0.1	-0.424	0.617	0.895
0.2	-0.569	0.596	0.838
0.3	-0.435	0.549	0.766
0.4	-0.357	0.404	0.504
0.5	0.019	0.258	0.280
0.6	-0.021	0.082	0.163

(4) 次数分布の異なる LFR ネットワークでの順位相関

8本から23本まで、ネットワーク内のノードの持つ最大次数が異なる16個のLFRネットワークを作成し、同様にして順位相関を比較した(図4.9)。各ネットワークにおいて、ノード数は100、コミュニティ数は20、平均次数は8であり、リワイヤリング率は0.1とした。ここで最大次数、平均次数が共に8の時、全てのノードの次数は8であるが、平均次数が8のまま最大次数を増やすと、その分、次数分布に偏りが出る。すなわちノードの持つ次数の差が大きくなり、次数の多いハブと次数の少ないノードの格差が大きいスケールフリーネットワークに近づくことを意味する。こうした次数格差がネットワークの経路長にもたらす影響を分析するため、比較指標としてZ値、P値、PWの値に加え、前節においてネットワーク可視化に使用した、他の3つのノード指標である媒介中心性、Burtの拘束度の逆数、Katz中心性を追加し、前節と同様の相関分析を行った。

16の全てのネットワークにおいて媒介中心性が最も標準順位との順位相関が高く、他の指標は次数に応じて順位が異なる結果となった。最大次数が8-10の次数格差の小さなネットワークにおいては、PW値が媒介中心性の次に標準順位との相関が高かったが、最大次数が大きくなるほどBurtの拘束度の逆数およびZ値の相関が高くなり、PW値やP値と順位が逆転する現象が見られた。P値はPW値と似たような挙動を示したが、PW値よりは常に順位相関が低かった。この結果は、ネットワークの最短平均経路長に対する各ノードの寄与をノード指標に基づいて分析する際、次数の影響を考慮する必要があることを示している。

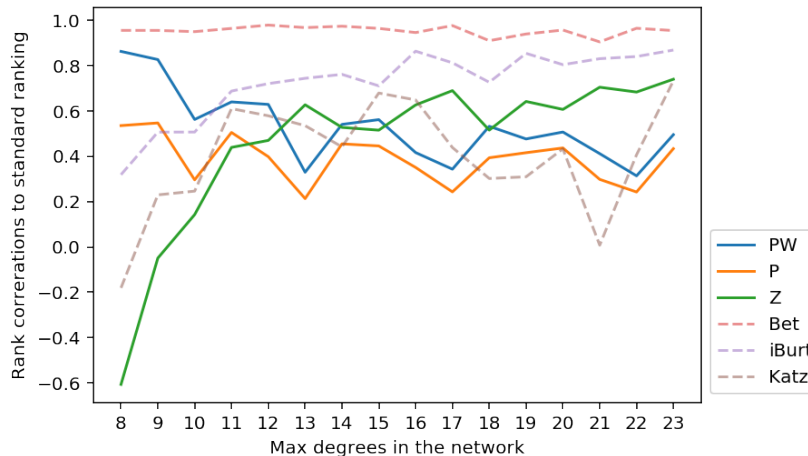


図 4.9: 最大次数の異なる LFR ネットワークでの順位相関. Y 軸は標準順位と比較した時の Spearman の順位相関係数, X 軸は各ネットワーク内のノードの持つ最大次数

(5) ノード削除とエッジ削除によるネットワーク径の変化

前節までの分析では, 各指標におけるランク順にノードを 1 つずつ削除していった際のネットワーク径 (最短平均経路長) を計測することで, 標準順位との相関を求めたが, 本節では, ノード削除による累積的な経路長への影響を分析すると同時に, 削除されたノードの次数の累積値にも着目する.

遠く離れたコミュニティをつないでいるノードの削除は, 近接したコミュニティをつなぐノード削除よりも, ネットワーク径を大きくすることが予想される. 遠いコミュニティをつなぐノードの不在は, 多くのネットワーク内経路が経由しなければならない重要なリンクを失わせる可能性があるからである. また, 攻撃耐性のコストは攻撃対象のノードが多くのリンクを持っているほど高くなることが予想される. たとえば, ソーシャルネットワークにおいて, 多くのフォロワーを持つインフルエンサーのように, 高い次数を持つノードはネットワークから削除したり制御することが難しい. 本節におけるノード削除実験は, 対象ノードを削除するために必要になるコストを考慮することに相当する (図 4.10).

ネットワークの連結性は, ネットワーク内の任意の 2 つのノード間の最短経路長の平均値 (ネットワーク径) によって示される. ネットワークの攻撃耐性に対処するために, Z 値, Katz 中心性, 媒介中心性, P 値, PW 値, iBurt 値, 次数中心性の 7 つのノード指標でランク付けされたノードを, 高順位ノードから連続的に削除した後のネットワーク径を東北ネットワークで測定した. 東北ネットワークは, 次数分布が 10 から 2000 までのスケールフリーな性質を持っている. 各ノードを削除した後のネットワーク径の大きさと各指標での順位を比較した. 比較のため, ノードをランダムに選択した場合のネットワーク径も同様に示している. 次数分布が不均質なネットワークに関する先行研究 [183] と同様に, 無作為選択の場合, ノード削除数が増加してもネットワーク径はあまり変化しない. 7 つのノード指標のいずれにおいても上位ランクのノードが除去されると, ネットワーク径は無作為選択の場合より速

く増大した。媒介中心性はノードを経由したリンク数を評価するため、ネットワーク径を短くするために最も影響力のあるノードを見つけるのに最も優れていたが、削除されたエッジ数を考慮すると、少ないエッジ数で最も影響力のあるノードの検出にはPWが最も優れており、PWが除去コストが低いノードを識別したことが示された。

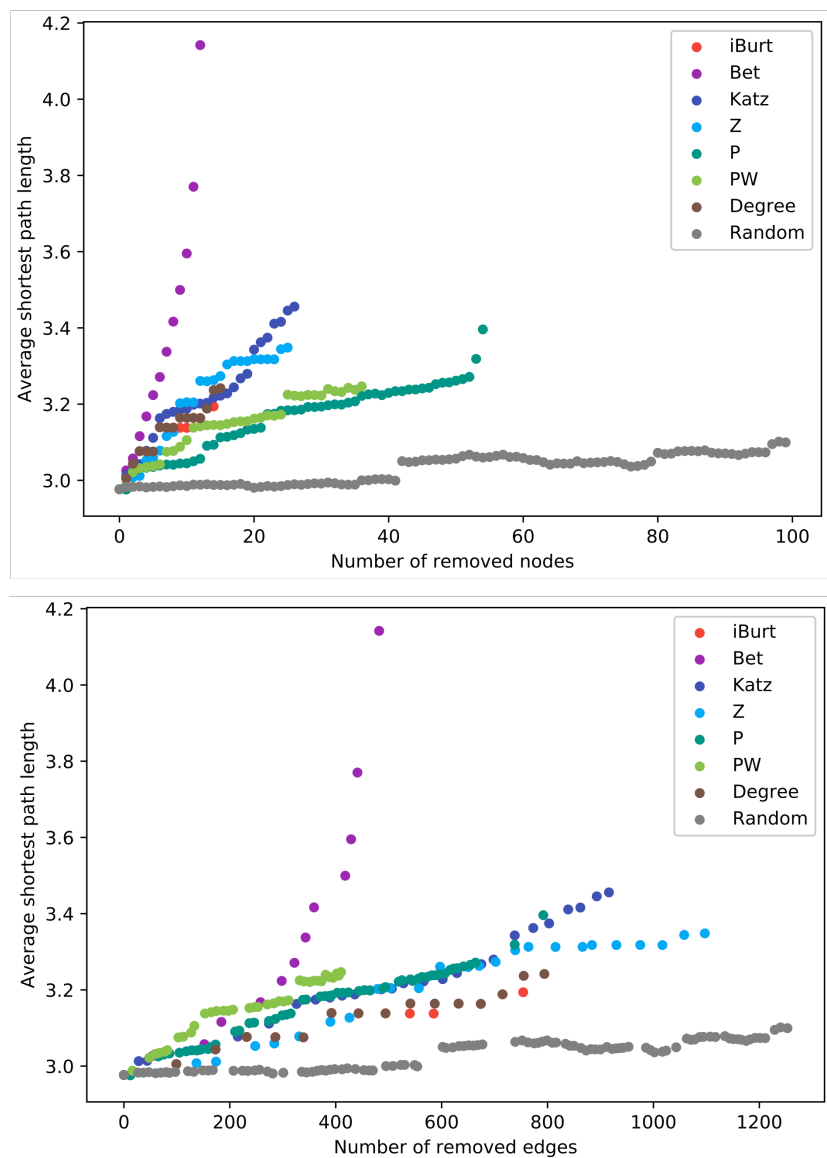


図 4.10: ノード攻撃後のネットワーク径の変化. Y 軸は攻撃後の東北ネットワークの平均最短経路長, X 軸は各攻撃で削除されたノードの積算数 (上図), ノード削除により失われたエッジ数 (下図) を指している

4.4 考察と結論

第4章では、不均一なコミュニティ構造を持つ複雑ネットワークの構造情報に基づく新たなノード指標PWを提案した。この指標はコミュニティを結びつけるノードの役割に着目したものであり、遠く離れたコミュニティをつなぐリンクを持つノードを高く評価する性質を持つことが、3種類のネットワークの可視化実験で確認された。ネットワークの攻撃耐性実験では、次数格差が大きいネットワークにおいては、提案指標が最も平均最短経路長を長くするノードを捉えていたが、次数分布の偏ったスケールフリー性の高いネットワークではハブを識別するZ値をはじめ、次数をノード評価に組み込んでいる指標の優位性が確認された。一方で、累積的ノード削除実験では、ノードと同時に削除されるリンク数に着目した場合、提案指標PWが最も少ないリンク数でネットワーク径を短くするノードを識別することがわかった。リンク数を削除コストと考えた場合には、最も小さなコストでネットワーク径を小さくするノードを識別したと捉えることができる。

以上のような不均一なコミュニティ構造を持つノードの役割に着目した分析は、多様な産業集積構造を有する地域ネットワーク上に存在する企業が、どのような役割を果たしているのかを分析する際に適用することが可能である。地域産業クラスターの中心に位置するハブ企業以外にも、クラスター間を結びつけるコネクタ企業、なかでも遠く離れたクラスター同士を結びつけるいわばブローカーの役割を果たす企業は、新たな事業ポテンシャルや新市場の探索が必要とされるこれからの社会において益々注目されることが予想される。

第5章 実データの適用によるクラスター進化 とブローカーの役割の特定

第2章で述べたように、企業間取引ネットワークは、自然災害や火災事故などの予期せぬ外部危機によって大規模な範囲に渡り寸断したり、新しいリンクに置き換えられることが知られている。しかし、ネットワークのメゾレベルの構造であるクラスターのダイナミクス、特に外部環境の変動に対するクラスターの耐性や適応力に関しては不明なことが多い。ネットワーク科学分野ではクラスター進化イベントを識別する多くの分析手法が提案されているが[158, 162, 161]、ネットワークの中でどのようなポジション価値を有するノードが進化イベントに寄与するのかは明らかにはなっていない。

本章では第3章および第4章で提案したノード指標を用いて、企業間取引ネットワークにおけるクラスターの進化に着目した可視化と回帰分析を行う。具体的には、10年間の日本企業の企業間取引ネットワークの中での関係性が遠いクラスター同士をつなぐ企業（ブローカー）を提案指標PWを用いて検出し、その動態解析を行った。その結果、ブローカーは取引数が少なく、入れ替わりが多く、クラスター間の移動性が高い傾向があった。また、提案指標を含む複数のノード指標を用いて統計解析を行った結果、クラスター進化に最も大きく寄与するのはブローカーであることが判明した。さらに、ケーススタディとして、クラスター進化の観点から特徴的な3つのクラスターパスを抽出して分析するとともに、ブローカーの果たす役割について事例に基づいた考察を行った[184]。

5.1 指標の定義と分析手法

本章の企業間取引ネットワークに関する分析は、以下の4つの部分から構成される。第一に、クラスターの安定性、第二に、ネットワークの推移および構造情報に関するノード指標を用いた企業分析、第三に、クラスターパスとブローカーの役割の観点からのケーススタディ、第四に、ノードのポジション価値を表すネットワーク指標を用いたクラスター進化の予測分析である。

5.1.1 クラスター進化に関する分析指標の定義

(1) 継承クラスター

第3章で述べた通り、企業間取引の新陳代謝度は大半の企業において0.1~0.2程度に限定されているため、企業間取引ネットワークのクラスターの動態は比較的安定しており、前

年クラスターと後年クラスターの関係进行分析すれば、継承されるクラスターを特定することができる。ただし、クラスターのエッジ構成は毎年変化するので、ある年に存在するクラスターが翌年どのクラスターに継承されたかは自明ではなく、計算によって求めることが必要となる。本研究ではある年のクラスターと最も類似度の高い翌年のクラスターを「継承クラスター」とし、その特定を以下の手順で行った。

まず、高速かつ高精度なモジュラリティベースのクラスタリング手法 [152, 153] であり、先行研究で広く用いられている Louvain 法 [185] を各年の企業間取引ネットワークにそれぞれ適用して、クラスターを取得しクラスター ID を付与した。なお、Louvain 法では、クラスターサイズを決定する resolution パラメーターを導入する拡張手法が提案されているが [186]、今回用いるのは重みなしの無向グラフであるため、resolution はデフォルト値の 1 とした。次に、クラスター進化を定量的に識別する Palla らの手法 [158] に従い、連続した 2 年間における前年クラスターと後年のクラスターの全ての組み合わせ（クラスターペア）における類似度を Jaccard 係数を用いて計算し、前年クラスターと最も類似度の高い後年クラスターを継承クラスターとした。2007 年から 2016 年までの全クラスターペアのうち、類似度上位 3 番までの Jaccard 係数は補足資料の表 A.12 に記載している。

(2) クラスターパス

10 年間に存在した合計 114 個のクラスターについて、上記の操作を繰り返し、最も Jaccard 係数の高いクラスターペアの系譜をたどることで継承クラスターの ID リストを取得し、これを「クラスターパス」とした。クラスターパスの一覧を表 5.1 に示す。

表 5.1: クラスターパス一覧（数字は各年における継承クラスター ID）

クラスターパス	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
1	0	1	1	2	1	0	1	0	0	0
2	1	0	0	1	2	1	2	1	1	9
3	*	*	4	5	5	10	0	5	6	2
4	2	6	2	3	3	2	3	2	2	1
5	3	4	9	10	8	7	8	9	9	10
-	4	6	2	3	3	2	3	2	2	1
6	5	7	8	9	7	6	7	8	8	5
-	6	4	9	10	8	7	8	9	9	10
-	7	9	10	3	3	2	3	2	2	1
7	8	5	6	7	6	4	9	10	4	6
-	9	0	0	1	2	1	2	1	1	9
8	10	8	3	0	4	3	4	3	3	3
-	11	11	3	0	4	3	4	3	3	3

クラスターの合併等の進化イベントにより、途中の年次において継承クラスター ID に重

複が生じた場合は、2007年におけるクラスターIDが小さい方のクラスターパスを選択して番号を付与し、重複しているもう一方のパスには「-」を記載している。なお、2009年のクラスター4は、2008年の全てのクラスターペアの中で2008年のクラスター4との類似度が最大で Jaccard 係数 0.12 であるが、この値は2008年のクラスター4の継承クラスターである2009年のクラスター9の類似度より小さい。このため、2009年のクラスター4は2008年からのクラスターパスをつなぐことができず、継承クラスターとはならない。しかし、2009年のクラスター4は2010年以降は安定して継承クラスターが存在しているため、2008年と2007年に相当するクラスターパスには継承クラスターIDの代わりに「*」を付けた。なお、クラスターペア間の Jaccard 係数が 0.08 以上 0.3 以下の場合はそのパスを「分岐」と呼ぶことにし、クラスターパスとは区別した。本章ではケーススタディとして、典型的なクラスター進化パターンを示す3本のクラスターパス 1,2,3 について詳細な分析を行う。その他のクラスターパスについての分析結果は、補足資料の図 A.20, 表 A.13, 表 A.14, 表 A.15, 表 A.16, 表 A.17, 表 A.18 にまとめて掲載している。

(3) クラスターパスの安定性

前項で説明したクラスターパスの安定性を企業の所属クラスター推移に基づいて計算した。クラスターパス i の y 年における安定性 $S_{i,y}$ は、クラスターパス i の y 年におけるクラスターのユニーク企業数 $G_{i,y}$ を、クラスターパス i の10年間の総ユニーク企業数 G_i で割った値として式 5.1 で表される。

$$S_{i,y} = \frac{G_{i,y}}{G_i} \quad (5.1)$$

(4) 企業のクラスターパス滞在率

企業のクラスターパス滞在率 (within-cluster stability) を、対象企業の各年の所属クラスターIDを10年間分合わせたIDリストとクラスターパスとのコサイン類似度により計算した。計算にはクラスターIDのカテゴリ値を one-hot 行列に変換し、クラスターに継承クラスターが存在しない場合には、全要素が0である列を one-hot 行列に加えた。Z 値, P 値, PW 値, 新陳代謝度, 取引数の各指標における上位 200 企業および下位 200 企業を各年の企業間取引ネットワークを用いて取得し、それらの企業の所属クラスター推移を分析した。各年における上位あるいは下位企業のクラスターパス滞在率の平均値を、その年の上位あるいは下位平均クラスターパス滞在率とした (図 5.5)。

(5) 非存続年数

企業の所属クラスターIDリストを取得する際、対象企業がデータセット中に存在しない年には、クラスターIDは取得できない。そこで、10年間の設定期間中、所属クラスターID

が取得できなかった年数をカウントし、その値の各指標における上位 200 企業および下位 200 企業における合計値を非存続年数 (withdrawal) とした (図 5.6)。

(6) クラスタ進化マップ

企業間取引ネットワークにおける企業のポジション価値とクラスターパスの安定性との関係、そして誕生 (birth)、合併 (merging)、分裂 (splitting)、死亡 (death) といったクラスター進化 [158] との関係进行分析するため、10 年間に存在した合計 114 個のクラスターの継承関係とノード指標毎の評価値を可視化した「クラスター進化マップ」を作成した (図 5.7, 図 5.8, 図 5.9, 図 5.10, 図 A.16)。円の大きさはクラスターを構成する企業数を示しており、クラスターは年次毎に下から企業数の多い順に並べている。円に記された数字は該当年におけるクラスター ID, 円の緑色の濃さは P 値, PW 値, Z 値, 新陳代謝度の 4 つのノード指標それぞれにおける評価値のクラスター毎の平均値を表している。赤い四角で囲んだ数字は 3 つのクラスターパスの番号を示しており、矢印上の数字はクラスターペアの Jaccard 係数である。視認性向上のため、P 値, PW 値, Z 値のクラスター進化マップには、指標毎に特徴が見出される典型的なクラスターパス 1,2,3 のみをそれぞれ矢印で示したが、新陳代謝度のクラスター進化マップにはクラスターパスの分岐を含めて描画し、Jaccard 係数が 0.3 以上の場合には直線の矢印で、Jaccard 係数が 0.2 以上 0.3 未満の場合は点線の矢印で、Jaccard 係数が 0.08 以上 0.2 未満の場合は細い点線の矢印で、全てのクラスターパスを記載し、クラスターパス 1,2,3 は赤い矢印で記載した。

(7) 地域エントロピーと産業エントロピー

各クラスターを構成する企業群の地域分布と産業分布に偏りがあるかどうかを調べるため、情報理論で一般に用いられるシャノンによる定義式 (5.2)[187] に従ってエントロピーを計算した。ここで、 P_i は事象 $X = i$ となる確率であり、地域エントロピーの場合は、分析対象となるある年のあるクラスター (対象クラスター) を構成する全企業のうち、所在地が地域 i である企業の割合である。計算対象となる地域は対象クラスターの全構成企業の所在地である東北 6 県と東京都であり、対象クラスターの構成企業が 1 社も存在しない地域は計算対象外となる。地域エントロピーが低いクラスターは、特定地域の企業が集積する地域クラスターである傾向が強い。産業エントロピーの場合もこれと同様に、確率 P_i は対象クラスターの全構成企業のうち、産業分類が産業 i である企業の割合となる。計算対象となる産業は対象クラスターの全構成企業の属する産業である。産業エントロピーが低いクラスターは、特定産業の企業が集積する産業クラスターである傾向が強い。

$$H(X) = - \sum_i P_i \log P_i \quad (5.2)$$

(8) クラスターパスのレーダーチャート

クラスターパス毎の特徴量の違いを分析するため、レーダーチャートによる可視化を行った。用いた特徴量は新陳代謝度 (metabolism), Z 値, P 値, PW 値, 当期決算売上高 (sales, 単位: 千円), 産業エントロピー (industry), 地域エントロピー (region), クラスターを構成する企業数を意味するクラスターサイズ (size) の 8 つの特徴量であり, このうち前半 5 つは企業固有の特徴量, 後半 3 つはクラスター固有の特徴量である。企業固有の特徴量は全てクラスターパスの該当年次における継承クラスターを構成する全企業の特徴量について最小値 0, 最大値 100 になるように正規化した後に平均値を取得しており, 外れ値の多い特徴量についてはレーダーチャート内に収まるようにスケールングを行った。具体的には, 新陳代謝度は 4 倍, Z 値は 3 倍, 当期決算売上高は 130 倍に正規化後のスケールングを施した。クラスター固有の特徴量は正規化は行わず, 産業エントロピーおよび地域エントロピーは各 30 倍, クラスターサイズは 0.1 倍のスケールングのみを施した。データセット中の全ての各特徴量について同じ処理を施しているため, 同一クラスターパス内における特徴量毎の年次推移の比較やクラスターパス間でのレーダーチャートの形状比較は可能だが, 異なる特徴量間で値の大きさを比較することはできない (P 値と PW 値はスケールングは行っていないが, 正規化した時点で元データの分散が反映されなくなるため, 正規化後の両指標の値の大きさの比較は意味をなさない)。

本章では典型的なクラスター進化の特徴を示す 3 つのクラスターパス 1,2,3 について, レーダーチャートをそれぞれ 2 種類作成した (図 5.12, 図 5.13, 図 5.14)。1 つは同一クラスターパス内で特徴量の年次推移を示すものであり, クラスターパスの年毎の継承クラスターを構成する企業群の平均値を特徴量毎に算出して描画した。なお, 視認性向上のため, 描画する特徴量は 1 年おきとし, 2008, 2010, 2012, 2014, 2016 年の値を色別に掲載した (クラスターパス 3 は 2008 年の継承クラスターが存在しないため, 2010 年以降の値を掲載した)。もう一つは, クラスターパス間のレーダーチャートの形状を比較するものであり, クラスターパス毎に継承クラスターの平均特徴量の 10 年間もしくは 8 年間の平均値を描画した。その際, 分析対象となるクラスターパスの平均特徴量は赤線で描画し, 残りの 2 つのクラスターパスの平均特徴量は灰色の点線で描画した。なお, 全てのクラスターパスの正規化・スケールング処理済みの特徴量は補足資料の表 A.13 に記載した。クラスターパス 1,2,3 以外のクラスターパスのレーダーチャートは補足資料の図 A.20 にまとめて掲載した。

5.1.2 クラスター継承度の回帰分析

統計解析および機械学習モデルを用いた回帰分析を行い, クラスター継承度について, 予測精度の最も高いモデルにおける各説明変数の寄与度を求めた。統計解析では単回帰および多変量回帰分析に基づいて説明変数の係数や統計量の評価を行い, 機械学習モデルには比較的高い精度を出す回帰モデルとして汎用的に使用されている Elastic Net での正則化を適用した一般化線形回帰 (Generalized linear model: GLM-net), Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost) を選択し, 各説明変数の

重要度を比較した。

(1) 変数の定義と探索的データ解析

本節で目的変数に用いるクラスター継承度 Q_y^j は, Takaffoli らの提唱したクラスター進化検出手法 MODEC (Modeling and Detecting Evolutions of Communities) [161] に準ずるものである。連続する2年間で $y-1$ 年, y 年とした時, $y-1$ 年に存在したクラスターと y 年に存在したクラスターを組み合わせることで作られる, 全てのクラスターペアについて計算されるクラスター類似度であり, 式 5.3 で定義される。ここで, G_{y-1}^i および G_y^j は $y-1$ 年, y 年それぞれにおける対象クラスターを構成する企業の集合であり, 分子は G_{y-1}^i および G_y^j の積集合の大きさを, 分母は G_{y-1}^i と G_y^j のうち大きい方のクラスターの企業数を示している。

$$Q_y^j = \arg \max_{G_y^j \in G_y} \left\{ \frac{G_{y-1}^i \cap G_y^j}{\max(G_{y-1}^i, G_y^j)} \right\} \quad (5.3)$$

説明変数には提案指標である新陳代謝度, PW 値に加え, 代表的なネットワーク指標である次数, 次数中心性, 近接中心性, 媒介中心性, ページランク, Burt の拘束度, そしてコミュニティ構造を反映したノード指標である Z 値と P 値を使用した。また, 各変数の大まかなデータ傾向を把握するため, 下記の可視化による探索的データ解析を行った。

- 各変数のデータ分布の確認 (ヒストグラム)
- 各変数における外れ値の確認 (ボックスプロット)
- 変数間の相関分析 (相関関係プロット)

(2) 正規化および歪度補正

古典的な統計手法とは異なり, 機械学習の多くの手法では変数の正規性を要求しないものの, 正規性を持たせる, もしくは正規分布に近づけることにより予測モデルのパフォーマンスが向上することが知られている。また, データに歪度があると, 推定値の分散が過度に大きくなることもあるため, クラスター継承度の回帰分析では, 変数は全て標準化し, 歪度補正を行った。なお, 変数変換によく用いられる box-cox 変換は負の値に対しては適用できないため, 歪度補正には下記で定義される Yeo-Jhonson 変換 [188] を適用した。

$$\psi(\lambda, y) = \begin{cases} (y+1)^\lambda - 1 / \lambda & \text{if } \lambda \neq 0, y \geq 0 \\ \log(y+1) & \text{if } \lambda = 0, y \geq 0 \\ -(-y+1)^{2-\lambda} - 1 / (1-\lambda) & \text{if } \lambda \neq 2, y < 0 \\ -\log(-y+1) & \text{if } \lambda = 2, y < 0 \end{cases} \quad (5.4)$$

(3) 変数間の相関関係と変数選択

重回帰分析に用いる説明変数は9つのノード指標，すなわち，Z 値，P 値，PW 値，次数中心性 (Degree)，近接中心性 (Closeness)，媒介中心性 (Between)，ページランク (Pagerank)，Burt の拘束指標 (Burt)，取引数 (Neighbor) の中から，分散拡大係数 (VIF: Variance Inflation Factor) の基準に基づいて選択した。VIF は推定された回帰係数の分散が，多重共線性のためにどれだけ増加したかを測る指標である。VIF は1つの説明変数を目的変数とした際の，他の説明変数による重回帰分析での決定係数を用いて，下記で定義される。

$$VIF = \frac{1}{1 - R^2} \quad (5.5)$$

ここで R^2 はモデルの当てはまりの良さを表す決定係数であり，残差の二乗和を観測値の平均値 \bar{y} からの偏差の二乗和で割った値を1から引くことで求められる。 R^2 は観測値を $y = \{y_1, y_2, \dots, y_N\}$ ，回帰方程式による推定値を $f = \{f_1, f_2, \dots, f_N\}$ とする時，下記で定義される。

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - f_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (5.6)$$

(4) 機械学習モデルのアルゴリズム

線形回帰による回帰係数が高く，統計的有意性が得られた変数だとしても，非線形な機械学習モデルにおいて，同様に重要な変数とみなされるとは限らない。これは，非線形の学習モデルでは予測への寄与度を規定する指標が線形回帰とは異なることが多く，変数を異なる視点で評価しているからである。そこで，本章では，R の `parsnip` パッケージ [189] を用いて，Elastic Net による一般化線形回帰 (GLM-net)，RF，SVM，XGBoost を実装した。

一般に機械学習モデルで予測タスクを解く場合，変数間の相関関係は深刻な問題を招くことはないとされているが，今回は説明変数の予測への寄与度を知ることが目的であるため，線形回帰の場合と同様に，説明変数間の強い相関関係によって生じる多重共線性に注意する必要がある。今回使用した Elastic Net では正則化項により多重共線性の問題は回避されているが，RF の場合，それぞれの決定木の結果が相関性を持たないように，あえて説明変数をランダムに選択しているというアルゴリズム上の特性があるため，強く相関している変数が含まれていると，その特性が十分に機能しない組み合わせが発生する確率も高まる。以上を考慮し，強い相関関係がある説明変数はあらかじめ除外することとした。ここでは，重回帰分析の際に行った VIF に基づく変数選択を参照しつつ，RMSE に基づくモデル探索を行い，投入変数から次数中心性 (Degree)，ページランク (Pagerank)・媒介中心性 (Between) を除外した。

(5) ハイパーパラメーターチューニング

各アルゴリズムで使用したハイパーパラメーターの範囲を表 5.2 に示す。最適なパラメーター値の組み合わせ探索には理論上、各パラメーター値に対して全ての組み合わせを検討する必要があるが、ハイパーパラメーター数×100 通りをランダムに選択し、交差検証を行うことで妥当な結果が得られることが経験的に知られている。本章でもこの基準を採択し、各モデルについてパラメーター数×100 通りのランダム選択による Grid search を行った。すなわち、各モデルについて表 5.2 の範囲においてハイパーパラメーターのチューニングを行い、最もデータに適合したモデルを最終モデルとして採用した。

表 5.2: クラスター継承度予測の機械学習モデルにおけるパラメーター範囲

Name	Description	Range
Elastic Net		
penalty	モデルの正則化の総量	$[10^{-10}, 10^0]$
mixture	正則化項における Lasso と ridge の割合	$[0, 1]$
XGBoost		
trees	アンサンブルに含まれる決定木の数	$[50]$
mtry	決定木を分岐させるために使う説明変数	$[1, 7]$
min_n	スプリット後のノードに含まれる最低データ数	$[2, 40]$
tree_depth	決定木の深さの最大値 (スプリット数)	$[1, 15]$
learn_rate	アルゴリズムの学習率	$[10^{-10}, 10^{-1}]$
RF		
mtry	決定木を分岐させるために使う説明変数	$[1, 7]$
trees	アンサンブルに含まれる決定木の数	$[50]$
min_n	スプリット後のノードに含まれる最低データ数	$[2, 40]$
SVM		
cost	誤判別の許容度	$[2^{-10}, 2^{-1}]$
rbf_sigma	radial kernel の関数形を規定	$[10^{-10}, 10^0]$
margin	不感損失関数 (ϵ より大きければ残差)	$[0, 0.2]$

(6) 回帰分析の評価指標

機械学習モデルの評価指標は式 (5.6) で定義される決定係数 R^2 に加え、平均二乗誤差平方根 Root Mean Squared Error (RMSE)、および平均絶対誤差 Mean Absolute Error (MAE) を用いた。RMSE と MAE はどちらも平均化された誤差の大きさを表し、それぞれ下記の式 (5.7)、式 (5.8) で定義される。

$$RMSE = \sqrt{\frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}} \quad (5.7)$$

$$MAE = \frac{\sum_i^N |y_i - \hat{y}_i|}{N} \quad (5.8)$$

ここで N はサンプル数であり、 y_i は観測値、 \hat{y}_i は予測値である。MAE は誤差の分布にラプラス分布を仮定した場合の最尤推定と一致する。ラプラス分布は RMSE の誤差の分布に仮定される正規分布より裾野が広く、外れ値の影響を受けにくい。このため外れ値が多いデータの誤差を評価する場合には RMSE より MAE が適した指標といえる。

説明変数の重要度は、上記評価指標を用いて最も優れたパフォーマンスを示したモデルをベストモデルとし、そのモデルにおける説明変数の重要度として算出する。ただし、機械学習モデルにおける説明変数の重要度は、各モデルによって適切とされる指標が異なる。たとえば線形回帰では t 値の絶対値が重要度として頻繁に用いられるが、RF や XGBoost などの決定木アルゴリズムではジニ係数やエントロピーを用いて重要度を算出することが多い。本章では、各機械学習モデルの特性に応じて変数の重要度を R で算出するための一般的なフレームワークである、Variable Importance Plots : vip パッケージ [190] を使用した。

本章のクラスター継承度の予測における最上位モデルである、RF における説明変数の重要度は、vip パッケージでは 2 つの方法で算出される。一つは決定木の分類を学習する際に計算される予測の改善度であり、もう一つは学習に使われなかったデータである out-of-bag (OOB) を利用した並べ替え手法 (permutation method) による評価である。決定木では回帰ないし分類の予測の改善度を最大化する説明変数が選ばれ、説明変数の重要度は、二乗誤差の合計ないし、ジニ係数やエントロピーのような分割点の選択に用いられる基準の減少に基づいて算出される。OOB データは学習済みの各決定木の誤差を評価するために使われる。OOB データをそのまま使った場合の誤差と、ある 1 つの説明変数をランダムにシャッフルした場合の誤差との差分が大きければ、その説明変数は重要とみなされる。この誤差の差を全ての決定木について平均化したものが、各説明変数の重要度となる。

5.2 検証実験

5.2.1 データセット

本章では株式会社東京商工リサーチ (TSR) が提供する、2007 年から 2016 年までの 10 年間における東北 6 県 (青森県、岩手県、宮城県、秋田県、山形県、福岡県) の企業間取引データを用いて提案指標の検証およびケーススタディを行った。TSR の企業間取引データには各企業が取引している主要仕入先・主要販売先がそれぞれ 20 社まで収録されているため、1 社につき最大 40 社までの取引先企業が紐付けられている。第 3 章で使用した企業間取引データよりも多くの取引先情報が登録されており、きめ細かな分析が可能になるとともに、データ範囲が 10 年間と 2 倍になったため、より長期的な経時変化を捉えることが可能になった。東北地方の企業間取引ネットワークは、2007 年から 2016 年の間で、取引先のどちらか一方が東北地方に所在する企業である企業間取引データを用いて作成しており、収集された企業間取引データの総量は約 100 万件であったが、本章では次数分布の極端な偏りを

避けるため、取引先企業数が2000以上である企業および8未満である企業は除外してネットワークを作成した（以下、東北ネットワークとする）。各年の企業数 (Firms) と取引先企業数 (Edges) を表 5.3 に示す。

表 5.3: 東北ネットワークのデータ概要

Year	Firms	Edges
2007	2,396	19,162
2008	3,884	32,483
2009	3,801	31,488
2010	3,807	31,659
2011	3,701	30,879
2012	3,725	31,041
2013	3,852	32,326
2014	3,946	33,284
2015	4,105	34,985
2016	4,557	39,359
Total	37,774	316,666

(1) 地域別，産業別平均特徴量

本章で用いるデータの地域別，産業別平均特徴量を下記に示す（表 5.4，表 5.5）。特徴量には本論文の提案指標である新陳代謝度 (Metabolism) と PW 値，および比較指標である Z 値と P 値に加え，売上高 (Sales, 単位：千円) と第 3 章で定義した存続年数 (Duration), 取引先企業数 (Edges) を加えた。表は地域別，産業別平均特徴量いずれも新陳代謝度の値でソートした順に表示している。新陳代謝度と存続年数は定義上，期間を通じて企業に固有の値であるが，その他の特徴量は年毎に各企業で値が異なるため，ここでは各企業の存続年数における平均値を用いている。TSR の業種コードは日本標準産業分類と同様「大分類・中分類・小分類・細分類」の 4 段階構成となっているが，本章では大分類のみを使用している。なお，紙面の都合上，産業別平均特徴量は 10 年間を通じて企業数が合計 100 社以上の産業に限定して表示した。

地域別平均特徴量を見ると，東京都は新陳代謝度が最も低く，存続年数が最も短く，Z 値，P 値，PW 値，売上高，取引先企業数が最も多かった。新陳代謝度の最も大きい地域は宮城県であった。同県は売上高が東京都の次に大きかったが，存続年数は東京都の次に短かく，Z 値，P 値，PW 値は中位であった。福島県は存続年数が最も長く，宮城県の次に新陳代謝度が高く，PW 値も東京都の次に大きかったが，売上高は中位であり，Z 値と取引先企業数は最も小さかった。

産業別平均特徴量を見ると，複合サービス事業の P 値，PW 値が最も低く，新陳代謝度も 2 番目に低かった。日本標準産業分類によると，複合サービス事業とは、「信用事業，保険事

業又は共済事業と併せて複数の大分類にわたる各種のサービスを提供する事業所であって、法的に事業の種類や範囲が決められている郵便局、農業協同組合等が分類される」とされている。新陳代謝度と存続年数以外の全ての特徴量が最も高いのは、電気・ガス・熱供給・水道業であった。同産業は新陳代謝度は低位であったが存続年数は2番目に長かった。新陳代謝度が最も高い産業は建設業であったが、同産業のその他の特徴量は中位であった。

表 5.4: 地域別平均特徴量

Industry	Metabolism	Z	P	PW	Sales	Duration	Edges
東京都	0.084	0.491	0.492	0.712	3.274e+08	6.734	23.490
岩手県	0.117	-0.118	0.455	0.656	2.150e+06	9.391	12.483
山形県	0.130	-0.041	0.389	0.600	2.0005e+06	9.447	13.466
秋田県	0.136	-0.115	0.420	0.599	1.792e+06	9.228	12.051
青森県	0.148	-0.108	0.438	0.613	4.524e+06	9.300	12.594
福島県	0.149	-0.135	0.443	0.673	3.319e+06	9.460	11.266
宮城県	0.154	-0.079	0.432	0.619	6.704e+06	9.124	12.995

表 5.5: 産業別平均特徴量

Industry	Metabolism	Z	P	PW	Sales	Duration	Edges
生活関連サービス業、娯楽業	0.090	-0.329	0.355	0.574	9.924e+06	8.931	9.122
複合サービス事業	0.096	-0.036	0.241	0.371	1.128e+07	9.100	11.413
電気・ガス・熱供給・水道業	0.102	0.566	0.530	0.879	5.989e+08	9.714	24.841
製造業	0.102	-0.186	0.392	0.591	7.458e+07	8.947	11.580
学術研究、専門技術サービス	0.103	-0.054	0.435	0.676	6.687e+07	8.851	11.420
卸売業、小売業	0.105	0.163	0.440	0.649	3.698e+07	8.917	16.961
鉱業、採石業、砂利採取業	0.105	-0.122	0.495	0.730	9.041e+06	8.969	12.097
不動産業、物品賃貸業	0.110	-0.164	0.531	0.787	3.919e+07	8.705	14.405
サービス業その他	0.111	-0.348	0.487	0.753	7.772e+06	9.052	9.533
情報通信業	0.115	0.125	0.337	0.561	8.266e+06	9.725	10.851
運輸業、郵便業	0.135	-0.197	0.407	0.660	3.018e+07	9.277	10.448
建設業	0.169	-0.084	0.446	0.627	7.454e+06	9.338	12.821

(2) 地域別、産業別企業数のクラスター分布

東北ネットワークのクラスターは、地域や産業によって特徴づけられ、クラスターを構成する企業数の分布傾向は、10年間を通じてほぼ同様であった（補足資料：図 A.18, 図 A.19）。以下では例として、2016年の東北ネットワークにおけるクラスター別企業数を地域毎に示す（図 5.1）。クラスター 2,3,4,9,10 は地域固有のクラスターであり、それぞれ宮城県、秋田県、青森県、山形県、岩手県に所在する企業が大半を占める地域クラスターであった。宮城県と福島県は他県に比べてクラスター間の企業分散が大きく、地域をまたいだ取引が行われている傾向が強かった。

図 5.2 は、2016年の東北ネットワークにおける企業のクラスター分布を産業ごとに階層化して示したものである。各セルの青色の濃さは産業別に標準化した企業数を反映しており、産業毎に見た場合の企業のクラスター分布を示すものとなっている。産業毎に企業数は大き

く異なるため、青色の濃さで産業間の比較はできない。分析の結果、10年間を通じて各クラスターに特徴的ないくつかの産業が存在することがわかった (A.18)。また、先述した地域固有のクラスターの大半は、伝統的な1つの産業を特徴とする傾向があることがわかった。例えば、クラスター9は山形県の建築業と農業・林業が主要産業であり、クラスター4は青森県の石材・砂利の採掘・採石業、クラスター10は岩手県の石材・砂利の採石業が主要産業であった。これに対して、主要産業が多岐にわたるクラスターはクラスター0,1,7であった。各クラスターを構成する産業はそれぞれ、クラスター0: 漁業/宿泊業, 飲食サービス業/複合サービス事業/製造業/卸売業, 小売業/農業, 林業, クラスター1: 運輸業, 郵便業/金融業, 保険業/サービス業その他/不動産業, 物品賃貸業, クラスター7: 学術研究, 専門・技術サービス業/情報通信業/生活関連サービス業, 娯楽業であった。

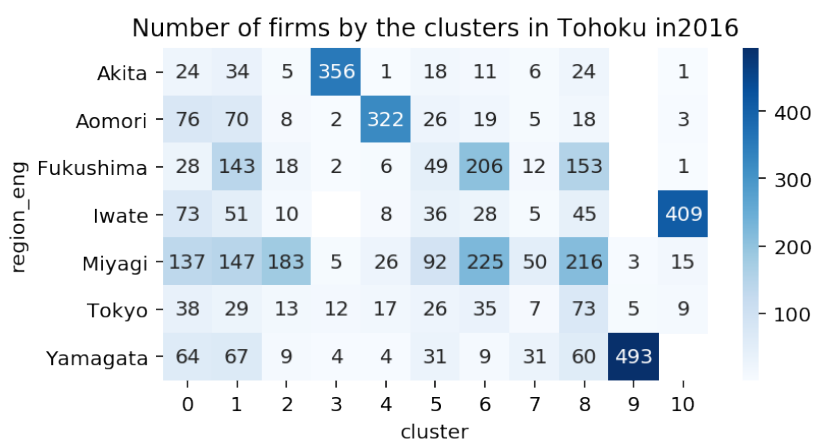


図 5.1: 2016 年の地域別・クラスター別企業数

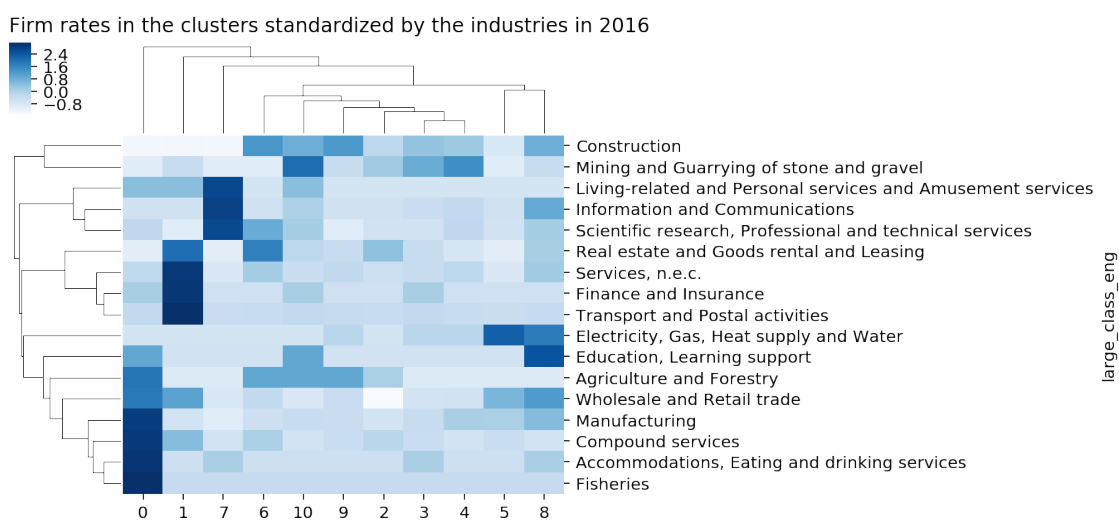


図 5.2: 2016 年の産業別・クラスター別企業割合

5.2.2 ブローカーの特徴

本節ではネットワークの構造的特徴を反映する既存指標である Z 値, P 値, そして第 4 章で提案したノード指標 PW 値に基づいて各年ごとに企業を評価し, 東北ネットワークの推移過程におけるそれぞれの特徴を比較検証する. 特に PW 値が高い企業をブローカーとし, 他指標で評価値の高い企業との特徴の違いに着目して分析を行った.

(1) 特徴 I - 希少な取引先 -

分析対象企業を Z, P, PW 値のスコアに基づいて上位 5%, 下位 5%, それ以外の企業の 3 段階に分け, それぞれ High, Middle, Low クラスとした. 先行研究 [47, 55] により, Z 値が高い企業はクラスター内の企業をつなぐ「ローカルハブ」, P 値が高い企業は異なるクラスターに所属する企業間をつなぐ「コネクター」としての役割を持つことが知られているが, PW 値が高い企業は, コネクターの中でも特に異質なコミュニティに属する企業をつなぐ役割があることから「ブローカー」とした. これら 3 つの役割および指標スコア別クラス分類毎にそれぞれの隣接ノード数, すなわち取引先の企業数を比較した.

企業の取引先数には指標スコア別クラス分類毎に共通した分布傾向が見られ, 3 つの指標のいずれにおいてもスコアの高い企業ほど, 隣接ノード数が多かったが, ローカルハブはコネクターやブローカーよりも隣接ノード数が顕著に多い傾向があった. また, ブローカーの方がコネクターよりも隣接ノード数がやや少なかった (図 5.3). 各指標のスコア別分類から見た次数相関は Z 値が最も強く, PW 値が最も小さい傾向があった. Z 値は次数の影響を大きく受けるが, PW 値は Z 値, P 値ほどは次数の影響を受けない傾向があることが伺える. これらは 10 年間の実験期間を通じて共通して見られる傾向であった (図 A.17).

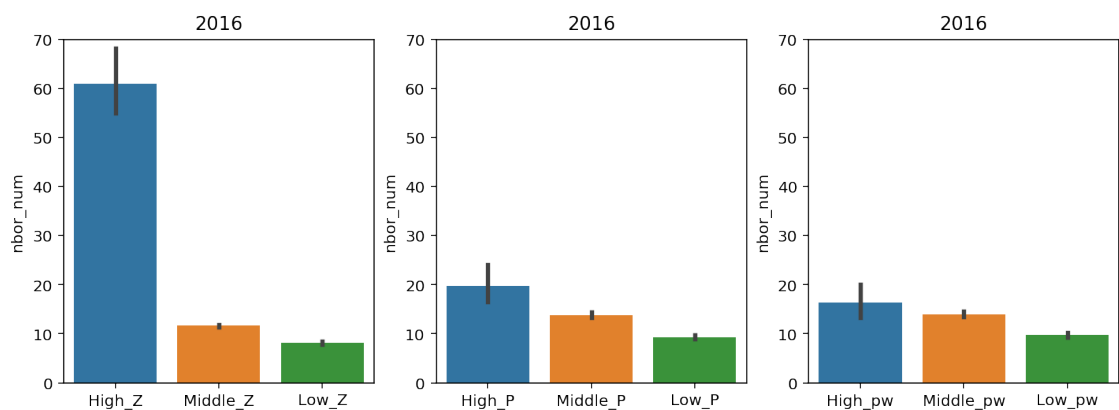


図 5.3: 2016 年における東北地方の企業の指標別取引数.

(2) 特徴 II - 高い入替率 -

前節において、ブローカーは相対的に隣接ノード数が少ない企業が多いことがわかった。本節ではネットワークの構造指標である Z 値, P 値, PW 値に, 隣接ノード数を加え, さらにネットワークの推移指標である新陳代謝度も含めて, 各指標で高く評価される企業集合の 10 年間の継続性を検証した。すなわち, 各指標スコアで上位 100 社に分類される企業集合を各年ごとに取得し, 連続する 2 年間における High スコア企業集合の Jaccard 係数を計算した。この作業を 10 年分繰り返し, 上位企業の推移を比較検証した。

5つの指標のうち, 隣接ノード数, すなわち, 取引先企業数が多い大企業に分類される企業集合が 10 年間を通じて最も安定しており, 次に安定な企業集合は Z 値の高いローカルハブであった。これに対し, P 値の高いコネクターや PW 値の高いブローカーは安定性が低く, ブローカーはコネクター以上に不安定であった。興味深いことに, 各指標毎に年次推移の変動幅を比較すると, 東日本大震災の前後である 2010-2011 年と 2011-2012 年の両時期には, 大企業に分類される企業集合は他期間より入れ替わりが大きい傾向にあるのに対し, Z 値, P 値, PW 値, 新陳代謝度の上位に分類される企業集合は, 他期間よりも類似度が高く, 相対的に企業の入れ替わりが少ない傾向があった (図 5.4)。こうした年次推移の変動幅は 5つの指標間の違いと比べれば僅かな差異ではあるが, 東日本大震災時のように広域的な企業間取引ネットワークの崩壊があった厳しいビジネス状況下において, Z 値, P 値, PW 値に反映されるネットワークの構造特徴, および新陳代謝度に反映されるネットワークの推移特徴が継承されやすい傾向があったことは注目に値する。

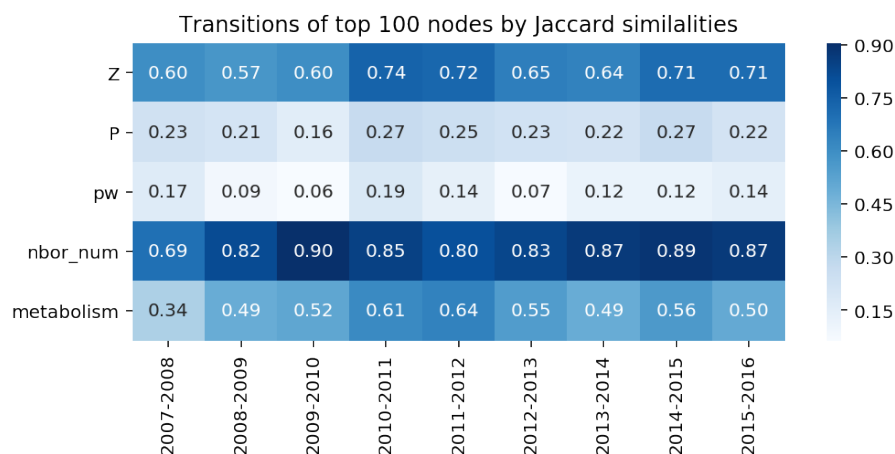


図 5.4: 上位 100 企業の指標別年次推移。

(3) 特徴 III - 高いモビリティ -

前節において, ブローカーに分類される企業集合は年次毎の入れ替わりが大きい, すなわち, ブローカーであり続ける企業は少ないことがわかった。そこで次に各年においてブローカーに分類された企業が 10 年を通じてどのようなクラスターパスを辿るのかを分析した。

所属クラスター滞在率が高ければ、その企業の辿るクラスターパスはクラスターパスに近くなるはずであるため、5.1 節で説明した方法により、各指標で上位 200 社に分類された企業群のクラスターパス滞在率 (within-cluster stability) を求め、指標間の違いを比較分析した。その結果、クラスターパス滞在率が最も高いのはローカルハブであり、2 番目は大企業であった。大企業のクラスター滞在率は無作為抽出とほぼ同じ水準であったが、新陳代謝度の高い企業は無作為抽出よりもクラスター滞在性が低く、コネクターやブローカーのクラスター滞在率はそれよりもはるかに低い水準であった。コネクターやブローカーはクラスターパスから離れて、クラスター間を移動しやすい傾向があることを示している (図 5.5 左)。

比較のため、各指標下位 200 社の企業群のクラスターパス滞在率を同様にして算出したところ、上位 200 社とは逆の傾向が見られ、P 値、PW 値でのスコアで下位 200 社の企業群 (以下、非コネクター、非ブローカーとする) のクラスターパス滞在率は、10 年を通じて他指標評価で下位の企業群よりも上位水準にあった (図 5.5 右)。

上位企業と下位企業の両者を比較すると、非コネクターおよび非ブローカーのクラスターパス滞在率は 10 年を通して 0.85 以上であり、ローカルハブや大企業よりもクラスター内に留まる傾向が強いことがわかった。また、Z 値、隣接ノード数、新陳代謝度では上位企業が下位企業より高いクラスター滞在率を示したのに対し、P 値、PW 値では上位企業が下位企業よりもクラスター滞在率が低かった。コネクターとブローカーは非コネクター、非ブローカーよりも頻繁にクラスター間を移動する、モビリティの高い企業であることが示された。

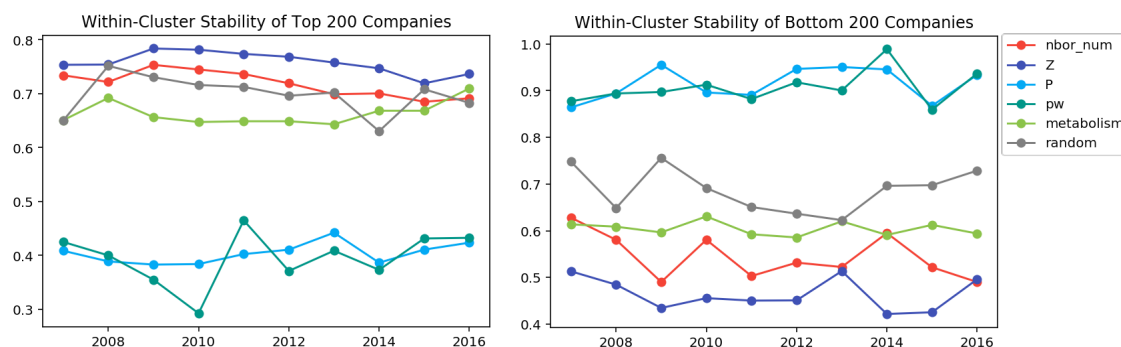


図 5.5: 上位 200 企業の指標別クラスター内安定性。

次に、5.1 節で説明した算出方法により、各指標における上位 200 企業の非存続年数 (with-drawal) を比較したところ、10 年を通じて新陳代謝度が上位の企業が最も非存続年数が長く (存続年数が短く)、コネクター、ブローカーは無作為抽出とほぼ変わらない水準であり、ローカルハブと大企業は存続年数が長い傾向があることがわかった。(図 5.6 左)。

比較のため、各指標下位 200 社の非存続年数を同様にして算出したところ、非ローカルハブ、非大企業は、ローカルハブ、大企業とは逆に存続年数が短い傾向があり、新陳代謝度の低い企業は、新陳代謝度の高い企業より存続年数が長い傾向にあった。一方、P 値、PW 値では上位企業と下位企業とで存続年数に大きな違いが見られず、いずれも無作為抽出と同水準であった。コネクター、ブローカーは存続年数とは相関がないことが伺える (図 5.6 右)。

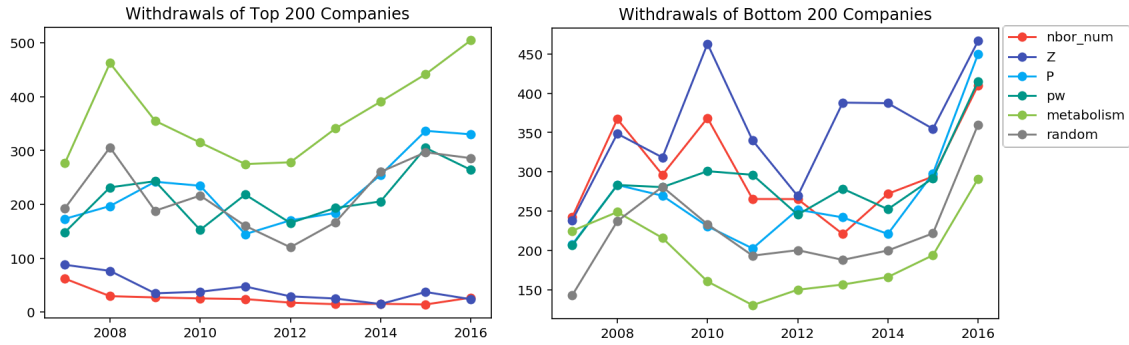


図 5.6: 上位 200 企業の指標別被消失数.

5.2.3 構造的ノード指標を用いたクラスター進化マップ

前節において、ブローカーに分類される企業集合は、取引先数が相対的に少ないこと、年次毎の企業の入れ替わりが大きいこと、クラスター間のモビリティが高いことがわかった。では、ブローカーのこうした性質は、所属するクラスターの進化にどのような影響を及ぼしているだろうか。第 2 章で見た通り、クラスター進化に関しては多くの先行研究があり [158, 144, 162, 161]、クラスターを構成するノードの評価をクラスターの進化イベントの予測に用いているものもあるが、その際、ノード指標の性質の違いを厳密には区別していないケースが多い [159, 191, 165]。そこで、本節では前節で行った 4 つのノード指標に基づく企業レベルの分析とクラスターレベルの分析を統合し、各年次の継承クラスターにおける企業のノード指標毎の役割分布とクラスター進化との関係を分析した。なお、クラスターパスの詳細な分析は 5.3 節で行うが、ここではクラスター進化マップの解釈に基づいて、継承クラスターの類似度とクラスター推移の関係に着目した分析を行う。

(1) クラスター外リンクを持たない企業の所在

はじめにクラスター間のモビリティに関して、ブローカーとは対照的な挙動を取ることが予想される、クラスター外リンクを持たず、クラスター内リンクのみをもつ企業の所在を確認した。その結果、大半のクラスターでは、構成企業の多くがクラスター内リンクとクラスター外リンクの双方を保持していたが、10 年間を通じてクラスター内リンクしか持たない企業がクラスターを構成する企業の半数ほどあるクラスターが 1 つ存在することがわかった。2016 年の東北ネットワークを例に挙げると、クラスター 0 ではクラスターを構成する全 450 企業のうち、半数近い 213 社の P 値および PW 値が 0 であり、クラスター内リンクしか保有していなかった (表 5.6)。

表 5.6: 2016 年の東北ネットワークにおけるクラスター内リンクのみを保有する企業数

Cluster ID	Within-Cluster Links Only	Within and Inter- Cluster Links
0	219	235
1	131	426
2	7	242
3	24	361
4	21	373
5	10	284
6	29	517
7	8	110
8	40	574
9	45	458
10	15	428

(2) コネクター (P 値) およびブローカー (PW 値) の分布推移

図 5.7, 図 5.8 は, それぞれ P 値および PW 値によるクラスター進化マップである。両マップを比較した結果, P 値と PW 値が最も低いクラスターについては, 両指標での評価が一致していたが, 評価値の高いクラスターでは両指標の評価が異なるクラスターが多数見られた。P 値が大きいコネクター企業は大きなクラスターにも見られるが, PW 値が大きいブローカー企業は比較的小さなクラスターに集中して存在する傾向があることがわかった。

最も安定したクラスターパスは 2007 年のクラスター 0 から始まるクラスターパス 1 であり, 連続する 2 年におけるクラスターペアの類似度が 10 年間の期間を通じて高く, P 値, PW 値は期間を通じて最も低い値であった。クラスターパス 1 は前節で述べたように, クラスター内リンクのみを保有する企業数が半数以上のクラスターで構成されていた (表 5.6)。クラスター外の企業と取引関係がない企業の割合がクラスターパスの安定性に寄与している可能性が高い。詳細は 5.3.2 節のケーススタディで説明するが, クラスターパス 1 内のクラスターは, 製造業や卸売業を中心とする, 地域や業種が多様な企業で構成されていた (図 5.1, 図 5.2)。クラスターパス 2 は 10 年を通して P 値, PW 値が中間的な値であり, 特に評価値や推移に特徴は見られなかった。2009 年のクラスター 4 から始まるクラスターパス 3 は小規模クラスターの中では珍しく安定したクラスターパスであり, 8 年を通じてクラスター類似度が大きく下がることなく, クラスターサイズも大きく変化せずに存続していた。クラスターパス 3 は P 値, PW 値ともに高い値であったが, PW 値については, 2012 年から 2013 年にかけてクラスター間の類似度がやや小さくなるとともに PW 値も減少していた。詳細は 5.3.4 で説明するが, クラスターパス 3 は宮城県の建設業を中心とする企業で構成されており, コネクターやブローカーを多く含むと考えられる。

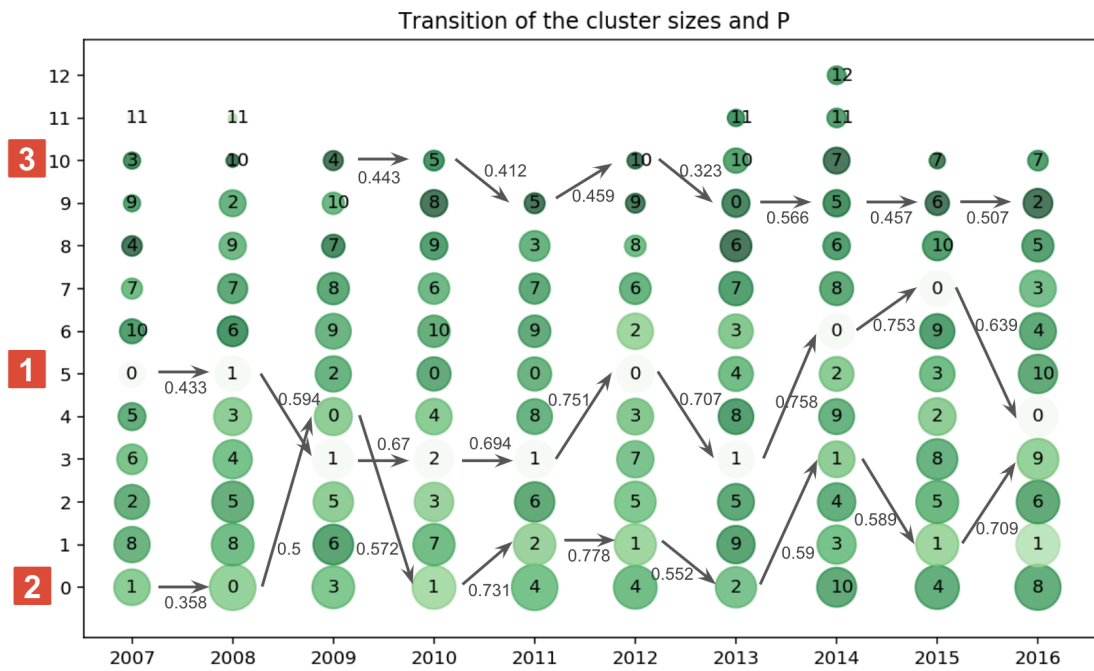


図 5.7: コネクターの分布とクラスターパス。

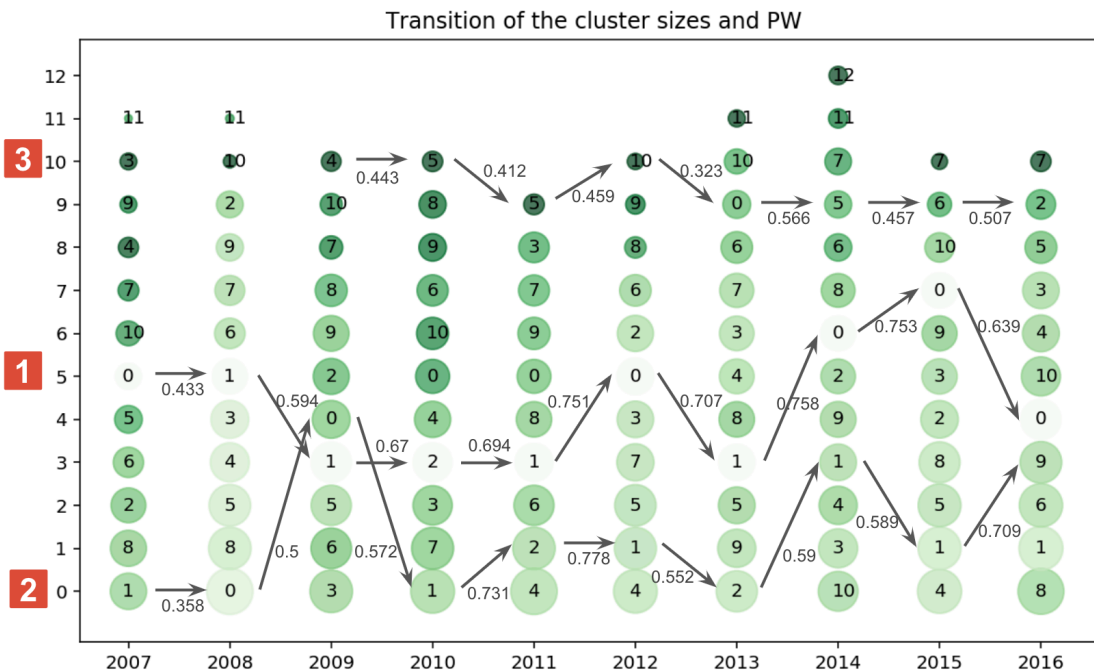


図 5.8: ブローカーの分布とクラスターパス。

なお、全企業の約10%に相当するPW値上位400社に限定して同様にクラスター進化マップを着色したところ、全企業で見たクラスター毎のPW値平均値（図5.8）とは異なる分布傾向が見られた（補足資料：図A.16）。特に、クラスターパス1は、クラスター内リンクしか持たない企業が多いためPW値が低くなっており、これらの企業が除外された場合は特別にPW値が低いわけではないことがわかった。

(3) ローカルハブの分布推移（Z値）

図5.9はクラスター進化マップをZ値で着色したものであり、前項のP値、PW値での評価とは対照的な傾向を示していた。すなわち、クラスターパス1はZ値は中間的な値であり、クラスターパス2はZ値が10年を通じて非常に高く、クラスターパス3はZ値が8年を通じて非常に低かった。クラスターパス2では、クラスターの安定性にローカルハブが重要な役割を果たしていると言えるが、クラスターパス3のように、ローカルハブが存在しなくても安定なクラスターパスもあることがわかった。詳細は5.3.3節のケーススタディで説明するが、クラスターパス2は、山形県の建設業、卸小売業の企業を中心に構成されている地域クラスターであった。クラスターパス2は、継承クラスター間の類似度が期間を通じて比較的高いにもかかわらず、合併や分裂といったクラスター進化イベントが多く見られることも特徴である（図5.10）。

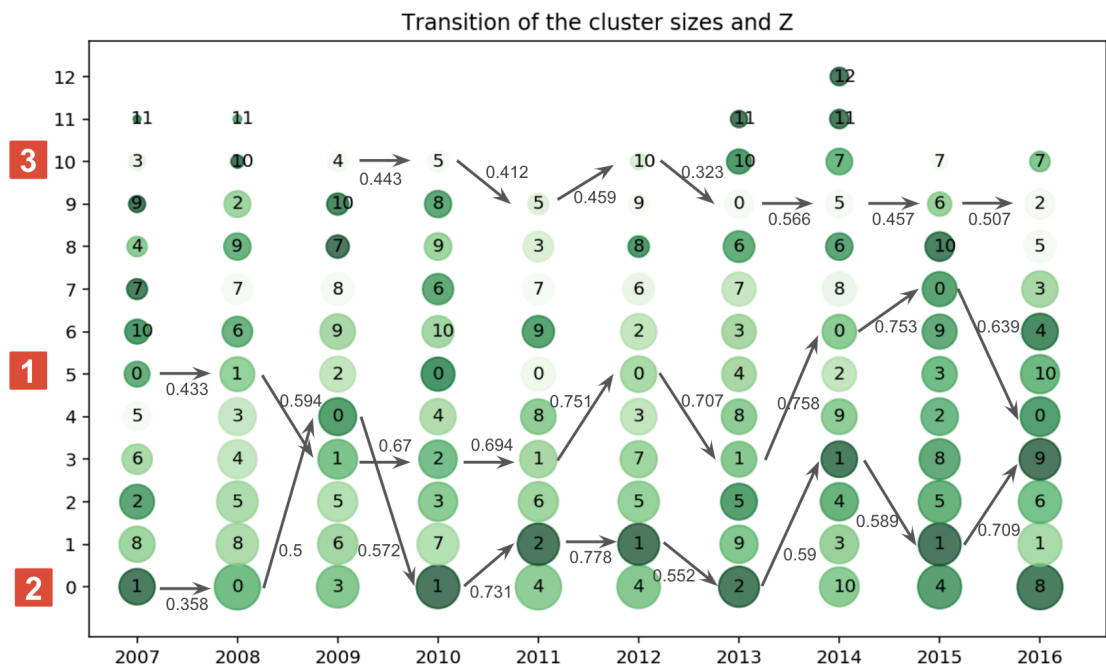


図 5.9: ローカルハブの分布とクラスターパス。

(4) 新陳代謝度の分布推移

図 5.10 はクラスター進化マップを新陳代謝度で着色し、全てのクラスターパスを分岐を含めて記載したものである。本節で着目するクラスターパス 1,2,3 は赤い矢印で示している。クラスターパス 2 は Z 値で見た場合と同様に新陳代謝度も中間値であり、特に大きな特徴は見られなかった。クラスターパス 1 と 3 については P 値および PW 値と新陳代謝度の評価値が似た傾向を示しており、クラスターパス 1 は新陳代謝度が期間を通じて低く、クラスターパス 3 は新陳代謝度が期間を通じて高い傾向があった。一方、クラスターパス 7 は、期間を通じて新陳代謝度が非常に高い傾向があったが、これは他の指標値には見られない特徴であった。クラスターパス 7 には宮城県と福島県の建設業と卸小売業の企業が多く含まれるが、地域エントロピー、産業エントロピーがともに比較的大きく、多様性の高いクラスターと言える（補足資料：図 A.20, 表 A.17）。なお、新陳代謝度は同じクラスターパス内では、評価値があまり変化しない傾向が見られたが、クラスターペアのどちらかもしくは両方の新陳代謝度が高く、クラスターペアの類似度が低い場合には、合併や分岐といったクラスター進化イベントが起きることが多かった。たとえば、2009 年のクラスター 10 は 2010 年のクラスター 3 に合併しており、2010 年のクラスター 3 は 2011 年のクラスター 3 とクラスター 9 に分裂している。

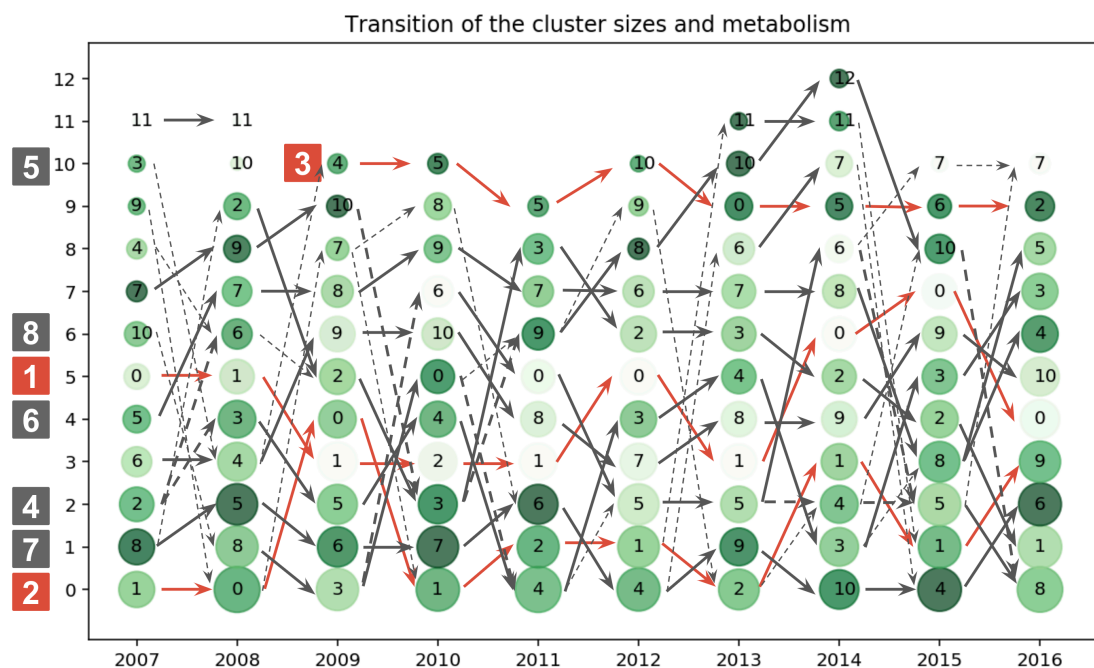


図 5.10: 新陳代謝度が高い企業の分布と分岐を含むクラスターパス。

5.3 クラスタースタビリティとブローカーのケーススタディ

本章では、企業間取引ネットワークネットワークにおけるクラスター進化と企業の構造的価値に関する典型的な遷移を抽出した。ここでは、上記の分析結果に基づき、3つの典型的なクラスター進化パターンを持つクラスターパスであるクラスターパス1,2,3について、そのパスに含まれる代表的な企業の特徴とともに記述する。

5.3.1 クラスタースタビリティ

2007年から2016年までの10年間における企業の所属クラスター推移に基づくクラスターパスの安定性を分析した(図5.11)。図中の凡例で示している番号は2007年における12個のクラスターのIDである。クラスターパスの安定性が大きく変化した場合には、クラスター同士が合併あるいは分裂するなど、クラスターのライフサイクルにおける進化イベントがあった可能性がある。2007年はデータ数が少ないこともあり、2008年以降の安定性とは大きく異なっている。2008年から2011年まではクラスターパスの安定性に大きなアップダウンが見られた。2012年以降はどのクラスターパスもそれぞれの安定性が大きく変動することは少なく、クラスターパスの安定性に関する順位は総じて維持される傾向があった。なかでも2007年のクラスターIDが0で始まる赤色のクラスターパスが最も安定しており、安定性は10年を通じて0.5程度を維持していた。これに対し、2007年のクラスターIDが2で始まる紫色のクラスターパスは安定性が0.2付近と最も低い水準であった。

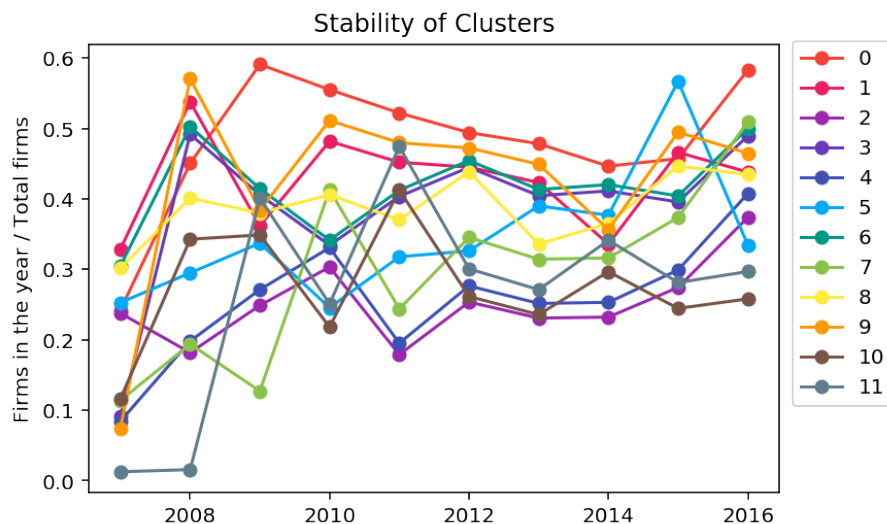


図 5.11: クラスタースタビリティ

5.3.2 地域多様性を持つ定常クラスター

全てのクラスターパスの中で最も安定性の高いクラスターパス1は、期間を通して約半数の企業がクラスター内リンクのみを持つ（表5.6）ことが特徴であり、誕生、合併、分裂、死亡といった大きなクラスター進化のイベント [158] は観測されていない（図5.10）。クラスターパス1は、全クラスターパスの中で新陳代謝度、P値、PW値が最も低く、地域エントロピーが高く、クラスター内に東北の全都道府県の企業が含まれていた。年次推移を見ると売上高は高く成長傾向がみられた（図5.12）。主要産業は卸売・小売業と製造業であり（表5.7）、そのほとんどがこの地域の典型的な食品産業である水産業に関連するものであった。

クラスターパス1の代表的な企業は、福島県を中心に隣接県への出店が奏功して強固な営業基盤を構築している卸小売業である。同社は生鮮食品をはじめとする、食料品販売を中心に店舗数を増やし業容は拡大基調にあり、10年間常にクラスターパス1に存在し、クラスターパスの安定性を維持する役割を果たしている。

また、Z値、PW値が共に高い企業を調べたところ、2007年の宮城県でZ値が9位以内でPW値が26位以内にランクインしている企業があった。この企業は卸売小売業の業務用食品会社であり、特に2011年の東日本大震災後も継続的に増収・増益・取引先数の増加を続けていた。また、同社は、長年にわたって多くのクラスター間のつながりがあったにもかかわらず、クラスターパスを離れることがなかった。

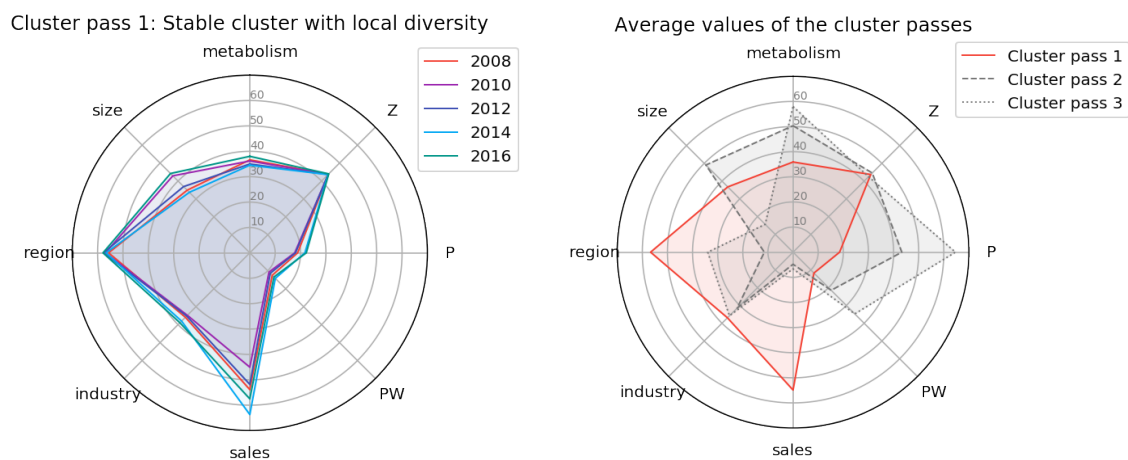


図 5.12: クラスターパス1の特徴推移とパス平均値の比較

表 5.7: クラスターパス 1 の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	0	1	1	2	1	0	1	0	0	0
卸売業, 小売業	94	151	201	197	189	163	154	140	145	187
建設業	3	4	3	4	5	5	5	5	3	4
漁業	3	3	3	3	3	3	3	2	3	3
生活関連サービス業, 娯楽業	19	24	2	3	2	1	1	1	1	1
製造業	59	153	197	186	169	177	168	159	166	201
運輸業, 郵便業	7	8	7	8	7	7	5	6	4	9
金融業, 保険業	2	2	2	2	1	1	1	1	1	2
宮城県	81	113	133	147	125	117	120	113	108	137
山形県	29	77	83	74	66	77	80	68	59	64
岩手県	11	26	66	52	55	51	39	43	49	73
東京都	18	27	33	34	29	32	33	29	29	38
福島県	5	28	42	41	32	22	27	25	25	28
秋田県	6	10	21	15	15	15	10	11	14	24
青森県	33	59	68	56	71	59	52	48	61	76

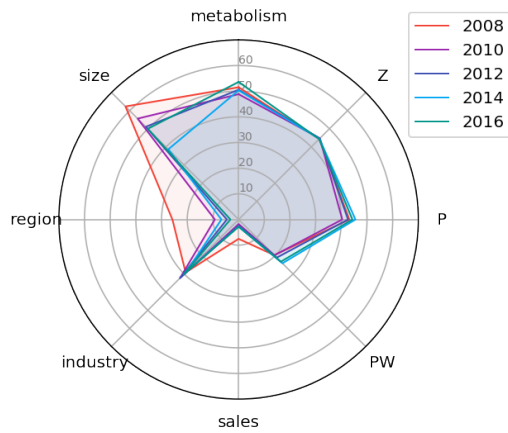
5.3.3 ローカルハブを中心とした地域クラスター

2番目に安定性の高いクラスターパスはクラスターパス2であり、Z値が高く地域エントロピーが非常に小さいことが特徴である。また、クラスターサイズは比較的大きいが売上高は非常に小さい(図 5.9, 図 5.13)。産業別・地域別企業数を見ると、クラスターパス2は山形県の建設業や卸売・小売業を中心とした産業で構成されていることがわかる(表 5.9)。ローカルハブであるZ値上位3社は、山形県内で50件以上の取引がある大手総合建設業であった。これらの企業は、県内の建設業を牽引する役割を果たしており、そのクラスター系譜はクラスターパスと一致していた。代表的な企業は山形県内を代表する地場トップクラスの有力ゼネコンである。同社は創業100年を超える老舗企業で、経営基盤は極めて盤石クラスターパス2に常駐する城主企業であり、進化イベントにつながるニッチを作り出している可能性が高い。クラスターパス1,3ではクラスター進化イベントが全く見られないのに対し、クラスターパス2ではクラスター進化イベントが多く検出されている。表 5.8 はクラスター進化マップ図 5.10 の分析により、クラスターパス2で検出された進化イベントの一覧である。表中の「分岐」はクラスターパス2の継承クラスターとの類似度が0.08以上0.2未満のクラスターのIDを、「進化イベント」は該当年に発生したクラスター進化イベントを示している。たとえば、2012年の継承クラスター1とクラスター9が合併して生じた2013年の継承クラスター2は、2014年に継承クラスター1とクラスター4に分岐している。

表 5.8: クラスターパス2の進化イベント

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	1	0	0	1	2	1	2	1	1	9
分岐	9		7			9	11	4, 11		
進化イベント		合併, 分裂		合併		分裂	分裂, 合併		合併	

Cluster pass 2: Regional cluster with local hubs



Average values of the cluster passes



図 5.13: クラスターパス 2 の特徴推移とパス平均値の比較

表 5.9: クラスターパス 2 の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	1	0	0	1	2	1	2	1	1	9
サービス業その他	8	4	8	15	13	12	10	8	20	13
不動産業, 物品賃貸業	1	3	5	9	9	5	7	7	10	7
卸売業, 小売業	69	118	59	94	89	82	83	61	84	78
建設業	218	434	302	375	351	349	347	280	364	354
製造業	27	44	26	47	40	46	28	22	43	35
複合サービス事業	2	1	2	1	1	1	2	1	1	2
農業, 林業	1	2	2	2	2	2	2	1	2	2
運輸業, 郵便業	40	8	7	10	12	7	5	5	7	6
宮城県	32	46	3	8	5	5	35	7	8	3
山形県	308	499	404	524	503	494	383	369	514	490
岩手県	4	9	-	2	-	-	11	2	1	-
東京都	4	15	3	6	4	5	15	2	3	5
福島県	6	23	1	4	-	-	23	-	1	-
秋田県	18	7	-	4	1	-	3	-	-	-
青森県	3	13	-	-	-	-	11	1	2	-

5.3.4 ブローカーが局在する小規模アジャイルクラスター

クラスターパス3は、P値、PW値、新陳代謝度が非常に高く、多くのブローカーが存在する、小規模かつアジャイルなクラスターで構成されていた(図5.14)。このクラスターパスは、クラスターペア間の類似度がクラスターパス1,2ほどは大きくはないが、8年間連続して継承クラスターが見られる、比較的安定したクラスターパスである。しかし、クラスターパス3はZ値が非常に低いため、その安定性はクラスターパス2で見られたような、ローカルハブによるものではないことがわかる。また、P値とPW値が非常に高いため、クラスターパス1で見られたような、クラスター内リンクしか持たない企業の存在によるものとも言えない。クラスターパス3は、宮城県の建設業、卸小売業、製造業を中心に構成されていた。しかし、他県や他産業の企業も一定程度存在しており、地域エントロピー、産業エントロピーともに中間的な値と言える。クラスターサイズは2012年に縮小したが、その後は成長する傾向が見られる。東日本大震災の影響を強く受けた、環境変動に敏感なクラスターパスであることがわかる。代表的な企業としては、土木、建築、漁港/港湾、舗装といった幅広い分野に業務展開している宮城県上位の総合建設業者があげられる。同社は新陳代謝度が非常に高く、東日本大震災関連の復旧工事・復興需要の増加で増収となり、損益は大幅増益を確保している。

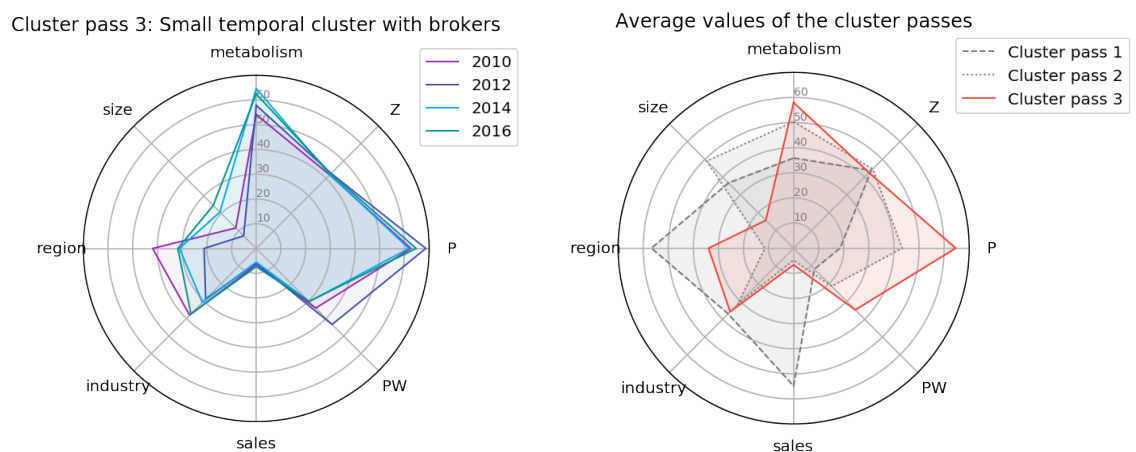


図 5.14: クラスターパス3の特徴推移とパス平均値の比較

表 5.10: クラスターパス 3 の産業別・地域別企業数

	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	4	5	5	10	0	5	6	2
サービス業その他	5	4	6	2	8	6	3	8
不動産業, 物品賃貸業	3	4	4	2	4	2	6	13
卸売業, 小売業	12	13	15	4	34	23	16	29
建設業	65	75	74	53	134	150	109	161
製造業	18	13	14	9	29	19	20	22
複合サービス事業	1	2	2	1	2	3	1	3
運輸業, 郵便業	3	4	3	1	5	2	5	5
宮城県	65	69	87	62	169	155	104	183
山形県	4	10	3	-	7	6	6	8
岩手県	9	7	2	-	7	5	13	10
東京都	8	8	8	5	10	11	11	13
福島県	14	15	11	4	11	18	15	18
秋田県	3	4	7	2	3	4	6	5
青森県	7	5	4	2	13	10	9	8

5.3.5 Role 分布のクラスターパス別年次推移

前節で典型的な 3 つのクラスターパスの特徴について述べたが、本節では、提案指標に基づく企業の役割（以下、Role とする）が、クラスターパスの特徴にどのように関係しているかを分析する。本節で着目する Role は、3 章で提案した新陳代謝度に基づく Changer と Holder, 4 章で提案した PW 指標に基づく Broker, 比較指標である Z 指標に基づく Hub, P 指標に基づく Connector, そして、保有する全てのエッジがクラスター内のみに存在する Inner の 6 つである。各 Role に相当する企業は各年に存在する企業のなかで、該当指標における上位 5% 企業を対象としている。Inner には指標が存在しないが、比較のため、各年における Inner から、その年の他の Role の企業数と同数をランダムサンプリングした。このため、全ての Role において、10 年間を通した該当企業数は 1888 社となっている。

図 5.15 は、各 Role に該当する企業の数、8 つのクラスターパス毎にそれぞれどのように年次推移したか（以下、Role 分布とする）を示している。まず、一見して Role 分布はクラスターパス毎に大きく異なることがわかった。すなわち、Broker はクラスターパス 3 に多く、Hub はクラスターパス 2 に、Inner はクラスターパス 1 に、Changer はクラスターパス 7 に多く存在していた。

該当企業の構成を見ると、クラスターパス 1 には Broker が比較的多く存在する一方で、Connector が少数であり、これらの Connector は全て Broker でもあることがわかった。これは、クラスターパス 1 には外リンクを持たない Inner が多いため、クラスターパス 1 のクラスターは他のクラスターと距離が遠くなっており、外リンクを持つ企業の多くが Broker になるためと考えられる。これに対し、クラスターパス 3 には Inner がほぼ存在せず、大半の企業が外リンクを保有している（表 5.6. 2016 年はクラスター 2 がクラスターパス 3 に相当する）。このため、クラスターパス 3 のクラスターは他のクラスターと距離が近い一方で、

遠くのクラスターをつなぐ企業も多いということがわかる。すなわち、クラスターパス3には Broker ではない Connector や、Connector ではない Broker が多数存在し、クラスターパス1の Broker や Connector 兼 Broker とは役割が異なることがわかった。

特徴的な Role 分布を示すクラスターパスとしては、他にクラスターパス7が挙げられる。クラスターパス7には Broker がほぼ存在せず、Changer が多数存在する。クラスターパス7は、平均新陳代謝度が最大であり、クラスター進化イベントも多い（図 A.20, 図 5.10）。クラスターパス7は、宮城県と福島県の建設業の企業を中心に構成されているが（表 A.17）、同じ建設業でもクラスターパス3とは Broker の局在が逆の傾向を示していた。各クラスターパスにおいて、Broker と Changer の増減には負の対応関係が見られ、Broker が増える年に Changer が減り、Broker の減る年に Changer が増える傾向があることがわかった。このことは、Broker の保有するエッジが切れにくく、Changer の保有するエッジが近くにあるクラスターをつないでいる可能性があることを示している。

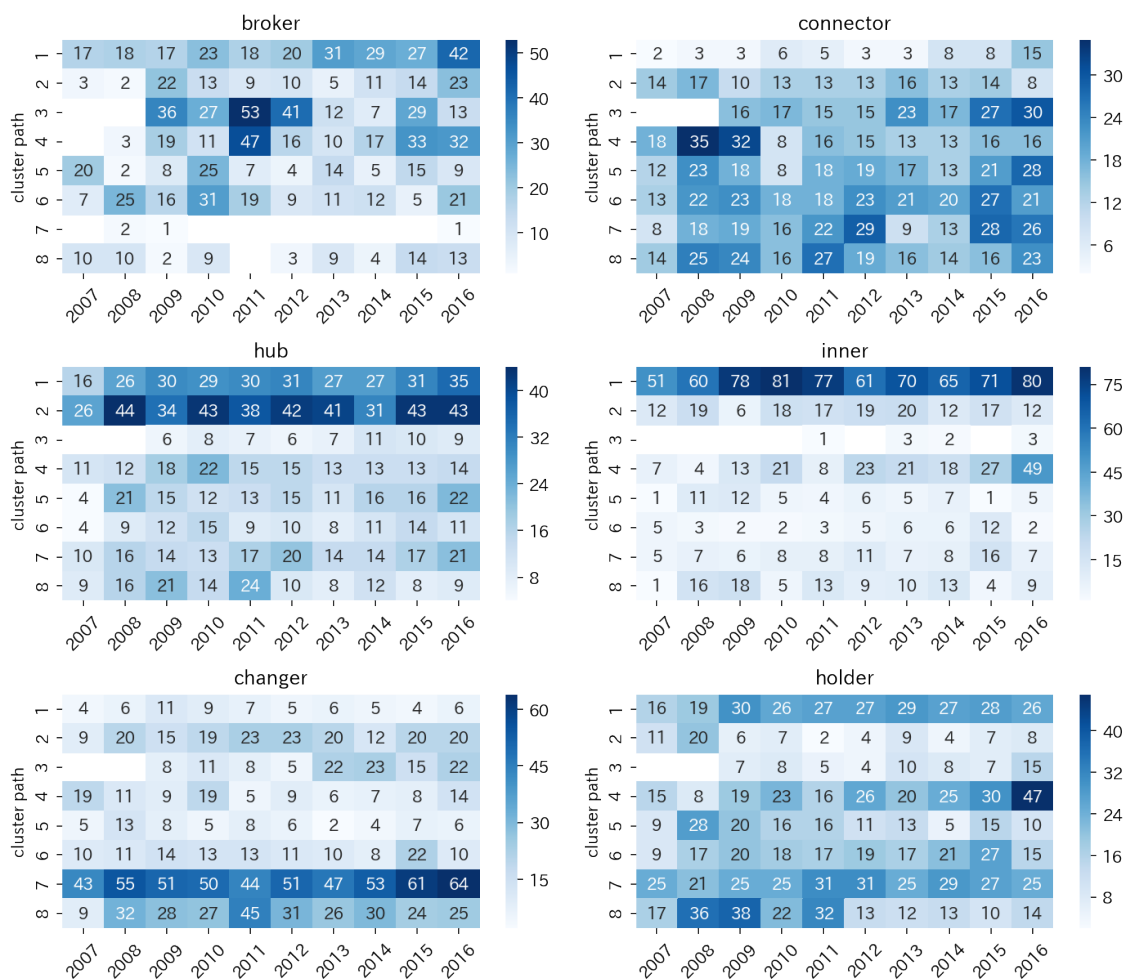


図 5.15: Role 毎に見たクラスターパス別企業数の年次推移

5.3.6 クラスターパス毎に異なるブローカーの役割

Role 分布の分析により、クラスターパス1とクラスターパス3には、いずれもブローカーが多数存在するが、その役割には違いがあること、新陳代謝度が最も大きいクラスターパス7にはブローカーがほぼ存在しないことが示された。5.2節では、ブローカーは入れ替わりが大きいこと、クラスター間の移動性が高いことを論じたが、本節では逆に、ブローカーである状態を維持する力が強いのは、どのようなクラスターパスであるかを分析する。

図5.16は、各クラスターパスに存在したブローカーが、そのクラスターパス内においてブローカーであり続けた年数（維持年数とする）の分布を示している。x軸は維持年数であり、y軸は維持年数の合計値を示している。すなわち、y軸の値をx軸の値で割った数は各クラスターパス内において、当該維持年数を持つブローカー数である。全てのクラスターパスにおいて、維持年数が1年であるブローカーが最も多く、維持年数が増えるに従って該当ブローカー数は減る傾向があった。また、クラスターパス1に存在するブローカーの維持年数が、他のクラスターパスに比べ、総じて長い傾向があることがわかった。

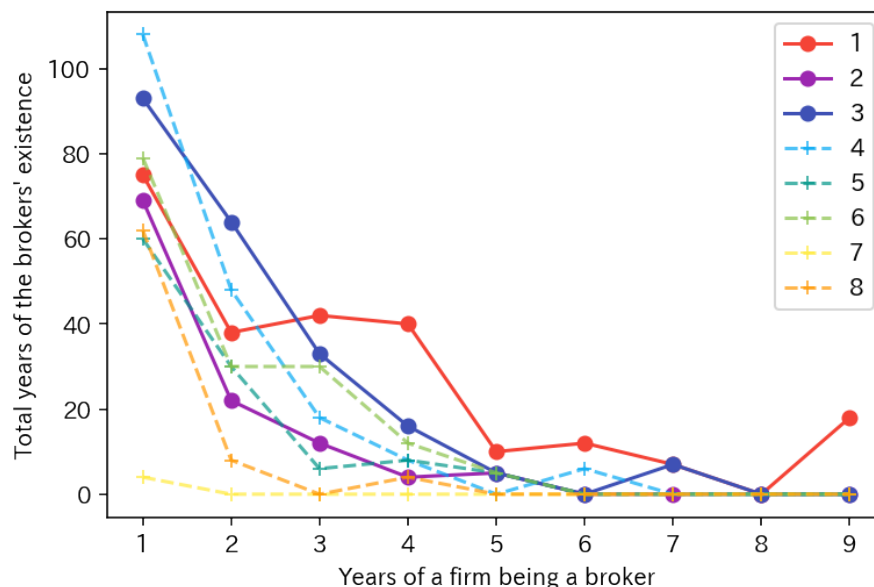


図 5.16: クラスターパス毎のブローカーの維持年数

5.3.7 ブローカーの維持年数と復興需要

前節で、ブローカーの維持年数はどのクラスターパスにおいても短く、大半のブローカーが1~2年でクラスターパスを外れたか、クラスターパス内にいたとしても、ブローカーではなくなったことがわかった。本節では個別のブローカーに着目し、ブローカーの維持年数と復興需要に関して、企業属性やクラスターパス間移動を踏まえた詳細な分析を行う。

(1) 維持年数が長いブローカーの特徴

本節では、クラスターパスを離れていてもブローカーであれば維持年数に含めることとし、全ブローカーの維持年数を調べた。その結果、10年間ブローカーであり続けた企業は3社であった。表5.11は維持年数7年以上の企業、合計15社の概要と平均評価値である。15社中、6社は東北地方ではない愛知県、東京都、大阪府に所在する企業であり、これらの企業は新陳代謝度が0であるところが大半であった。東北地方の中では宮城県の企業が6社と最も多かった。

表 5.11: ブローカーの維持年数と平均評価値

維持年数	地域	産業	取引先数	新陳代謝度	Z	PW
10	愛知県	卸売業, 小売業	14.78	0	-0.68	1.55
10	青森県	製造業	10.40	0	-0.57	1.43
10	宮城県	卸売業, 小売業	250.0	0.19	2.81	1.61
9	宮城県	サービス業(その他)	7.89	0.05	-0.96	1.35
9	東京都	卸売業, 小売業	9.11	0	-0.82	1.43
8	東京都	製造業	31.00	0	-0.01	1.22
8	大阪府	不動産業, 物品賃貸業	49.62	0.11	0.41	1.22
7	大阪府	不動産業, 物品賃貸業	19.29	0	-0.36	1.21
7	宮城県	建設業	8.57	0.16	-0.93	1.35
7	岩手県	建設業	9.29	0.05	-0.71	1.44
7	宮城県	学術研究, 専門・技術サービス業	10.40	0.10	-0.98	1.70
7	宮城県	製造業	12.00	0.03	-0.73	1.31
7	山形県	卸売業, 小売業	9.78	0.13	-0.67	1.24
7	大阪府	卸売業, 小売業	22.80	0	-0.35	1.23
7	宮城県	電気・ガス・熱供給・水道業	10.00	0.04	-0.42	1.17

維持年数が長い東北地方のブローカーに総じて言えることは、ハブ企業以外はいずれも、大手企業のグループ企業もしくは、官公庁ルートに強みを持つ企業が多いことである。取引先企業数が少なく、親会社を持たない中小企業等がブローカーであり続けるためには、何らかの支援が必要であることを示唆する結果と言える。また、維持年数の長いブローカーだとしても、1つのクラスターパス内に長期滞在するわけでは必ずしもないことがわかった。これはハブも例外ではなく、たとえば、維持年数10年の宮城県の卸売業・小売業は、常に250前後の取引先企業を有する城主企業だが、10年間に5回、異なる3つのクラスターパス間を行き来していた。新陳代謝度も比較的高い企業であることから、クラスターの変動やクラスター間のミキシングに影響を及ぼしている可能性がある。

(2) 復興需要の限界を乗り越えるブローカーの特徴

東日本大震災の復興需要により、東北地方の企業のなかでも瓦礫除去や土木工事等を行う建設業は大きく売上を伸ばした。しかし、政府による巨額の投資が落ち着くと売上伸長率が下がった企業も少なくない。本節では、復興需要の限界を乗り越えて売上を伸ばすブローカーの特徴に着目して分析を行う。

図 5.17 左図は、クラスターパス毎の平均値として見た当期売上伸長率の年次推移である。建設業の企業が多く滞在するクラスターパス 2, 3, 5, 6, 7 では、東日本大震災の翌年である 2012 年に当期売上伸長率が大きく伸びていた。図 5.17 右図は、クラスターパスに含まれる全企業を、当期売上伸長率の五分位点で分割し、各分位点に相当する企業の数でクラスターパス毎に比較したものである。卸小業・小売業を中心とするクラスターパス 1, 4 では、当期売上伸長率の中央値が含まれる分位点 40 に最も多くの企業が属しているが、建設業を中心とするクラスターパス 2, 3, 5, 7 では分位点 0 と 80 に多くの企業が属していることがわかった。建設業では、売上伸長率に関して企業間の格差が大きいことを示している。

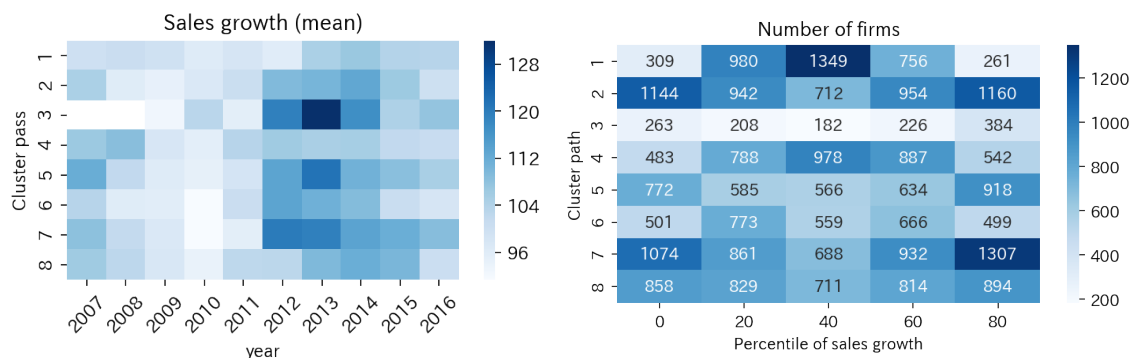


図 5.17: 売上伸長率のクラスターパス別年次推移と分位点分布

本節では、東日本大震災の翌年である 2012 年および、復興需要がある程度落ち着いていることが見込まれる 2016 年のどちらにおいても、当期売上伸長率が 110 より大きいことを、復興需要の限界を乗り越えて市場変動に適応した企業の条件（適応条件）とする。当期売上伸長率が 110 より大きいクラスターパス内の企業は 2012 年は 1080 社、2016 年は 822 社であった。このうち、適応条件を満たす企業は 216 社であり、その中で少なくとも 1 年はブローカーであった企業は 30 社であった。以下では、この 30 社のうち、対照的な当期売上伸長率の分布傾向を示した、クラスターパス 1, 3 に滞在したことがある企業に着目する。クラスターパス 1 に滞在した企業は 1 社、クラスターパス 3 に滞在した企業は 6 社であった。表 5.12 は、該当企業の概要と平均評価値である。維持年数は前節と同じく、ブローカーであった年数を示している。適応条件を満たすブローカーは全て建設業であり、7 社中、6 社が宮城県に所在していた。

表 5.12: 適応条件を満たすブローカーの維持年数と平均評価値

維持年数	地域	産業	取引先数	新陳代謝度	Z	PW	売上伸長率
4	宮城県	建設業	9.90	0.15	-0.52	1.15	133.70
1	宮城県	建設業	15.17	0.17	0.22	0.90	131.00
2	宮城県	建設業	10.43	0.42	-0.48	1.11	121.57
3	宮城県	建設業	24.00	0.15	0.62	1.06	115.62
3	宮城県	建設業	9.88	0.17	-0.24	0.80	104.62
5	岩手県	建設業	9.00	0.14	-0.64	1.19	103.62
1	宮城県	建設業	7.00	0.17	-0.52	1.06	94.80

適応条件を満たしたブローカーは、前節で分析した維持年数が長いブローカーに比べ、新陳代謝度が高く、PW 値が低い傾向があった。また、クラスターパス 3 では震災特需の影響が大きく、2016 年も復興事業を継続受注している企業が多いこと、このような企業はクラスターパス間の移動が少ないことがわかった。今回使用したデータは 2016 年が最新年であるが、政府による投資が落ち着いた際の変動に適応する企業については、より新しい年のデータを用いた分析が必要であると考えられる。

5.4 クラスター継承度の予測

前節までの分析により、第 3 章および第 4 章で提案したノード指標である新陳代謝度および PW 指標は、ネットワーク内のクラスター構造を識別する既存のノード指標である Z 値、P 値のいずれとも異なるクラスター進化への寄与度を示すことが推察された。すなわち、クラスター進化マップにより、ノード指標ごとに相関するクラスターパスが異なることが推察され、特にローカルハブとブローカーとではクラスター進化におけるノードの役割が大きく異なることが判明した。

クラスターがどのような経時変化を辿るのかについては、コミュニティイベント検出 [158] や MODEC [161, 165], GED [159] など多くの先行研究があるが、ノード指標毎に異なるクラスター進化への影響に着目した研究は、著者の認識する限りではまだ見られない。そこで本節では、クラスター進化の中でも特に重要な要素であるクラスターの継承度に着目し、回帰分析により各ノード指標のクラスター進化への寄与度を比較した。

5.4.1 探索的データ解析

分析に用いる 10 種類の連続変数について、探索的データ解析を行った。予測に用いる説明変数の候補となる変数は、媒介中心性 (betweenness), Burt の拘束度 (burt), 近接中心性 (closeness), 次数中心性 (degree), 取引数 (nbor_num), P 値 (P), ページランク (pagerank), PW 値 (PW), Z 値 (Z) の 9 つのノード指標であり、目的変数はクラスター継承度 (similarity) である。標準化後の各変数の分布状況をヒストグラム (図 5.18) により確認した。外れ値についてはボックスプロット (図 5.19) を用いた可視化により確認した。変数間の相関係数は相関関係プロット (図 5.20) によって確認した。ページランクと次数中心性など、変数間で相関関係が強い組み合わせが存在するため、変数選択に留意する必要があることがわかった。

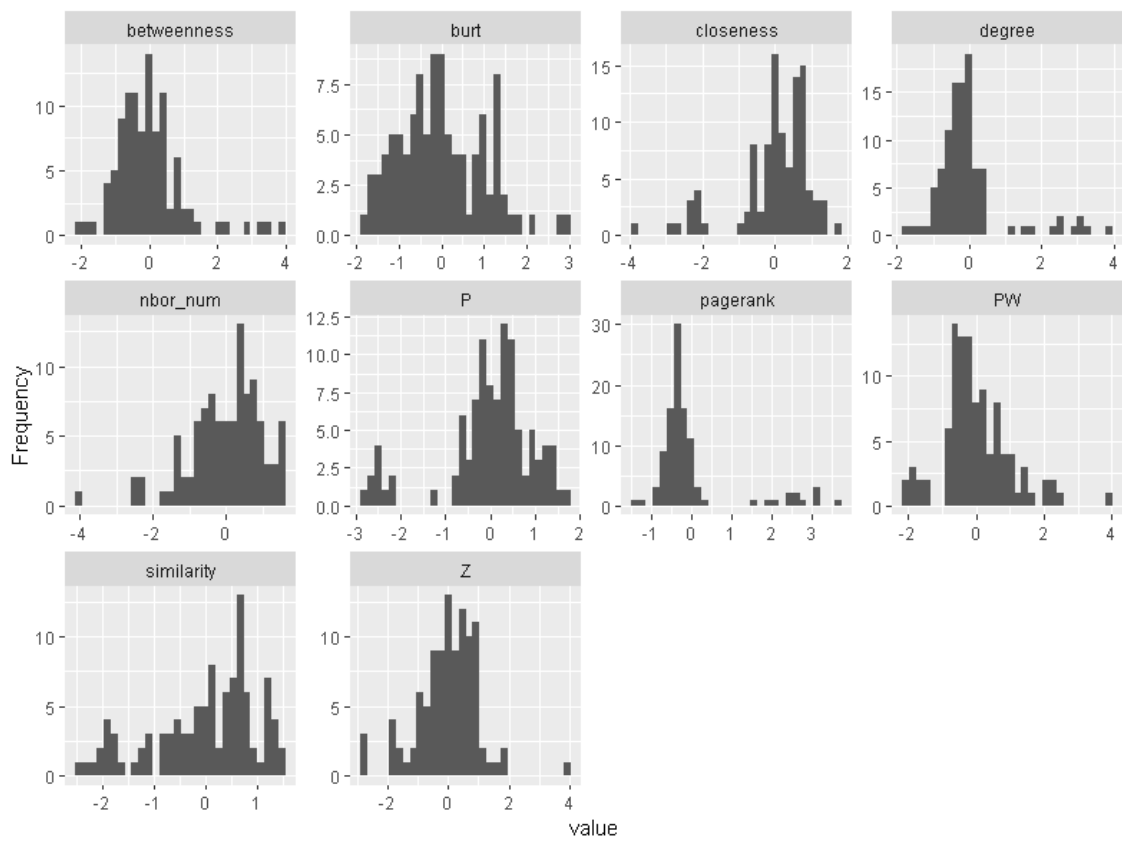


図 5.18: 各変数の値とデータ分布

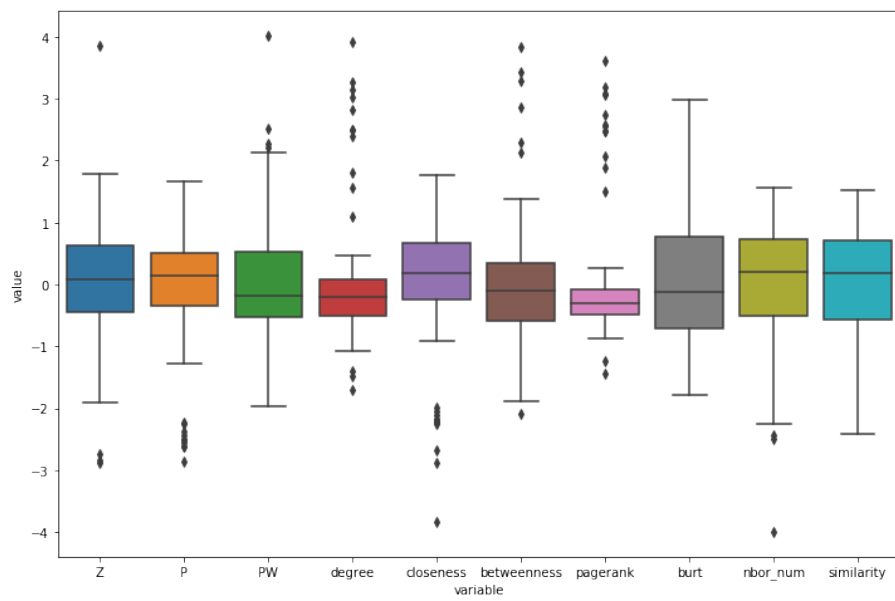


図 5.19: 各変数における外れ値の分布

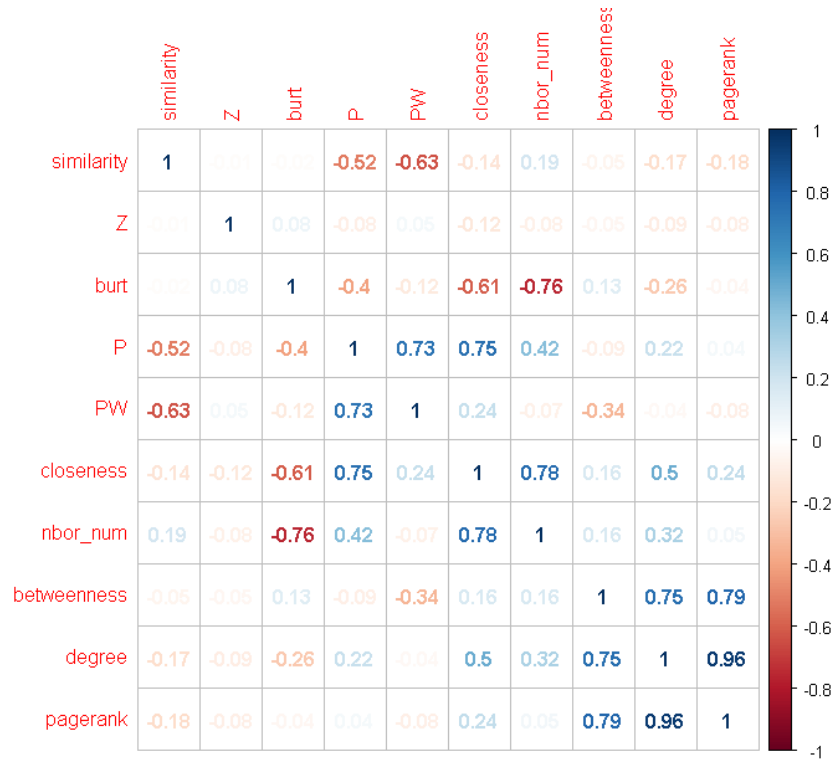


図 5.20: 変数間の相関関係

5.4.2 統計的回帰分析によるクラスター継承度の予測

(1) 単回帰分析

本節では最小二乗単回帰分析を用いて、前後年で最も類似度の高いクラスターペアの類似度を、前年のクラスターを構成する企業の各ノード指標における平均値から推定することにより、企業の役割とクラスター進化の相関を分析する。ノード指標は、Z 値、P 値、PW 値、新陳代謝度、および隣接ノード数を用いた。その結果、P 値と PW 値は p-value が有意水準を満たす 0.01 以下であり、PW 値は回帰係数 (Coef) が 5 つのノード指標の中で最も高い -0.511 であった (5.13)。この結果は、PW 値がクラスター継承度と弱い負の相関関係があり、PW 値が低いクラスターは安定的で、PW 値が高いクラスターは次年度へのクラスター継承が困難であることを示している。

定義上、PW 値が低いクラスターはクラスター内リンクが多く、PW 値が高いクラスターでは遠くのクラスターをつなぐクラスター外のリンクが多くなる。したがって、クラスター内リンクの割合が、クラスター構成要素の継承度と相関があると考えるのが妥当である。しかし、ローカルハブを検出する Z 値が統計的に有意水準を満たさなかった結果を考慮すると、クラスター内リンクを多く持つローカルハブの存在だけでは、クラスターの継承度を説明することはできない。また、P 値と PW 値における回帰係数の差は、クラスター間リンクの中でもより遠くのクラスター間を結ぶリンクが、クラスター継承度と強く相関する傾向

を示している。モデルのカバー率を表す R2 値、適合率を表す AIC は、単回帰分析の限界もあり、どの指標も十分に高い値ではないが、PW 値がそれぞれ 0.261, 260.075 と最もよい結果を示した。クラスター継承度の低さは、後年クラスターに合併、分裂、死といった大きな変動を伴うクラスター進化イベントが生じている可能性が高いことを示している。このため、クラスター継承度と PW に負の相関が見られたことは、ブローカーがクラスター進化において何らかの役割を果たしている可能性があることを示していると言える。

表 5.13: クラスター継承度とノード評価指標の回帰分析結果

Node Indicator	Coef	R2	p-value	AIC
Z	0.078	0.006	0.438	290.009
P	-0.38	0.145	0	274.84
PW	-0.511	0.261	0	260.075
Metabolism	0.027	0.001	0.791	290.554
Neighbor	0.287	0.082	0.004	281.95

(2) 重回帰分析

説明変数に主要なネットワーク指標を加えて重回帰分析を行った。前項の単回帰分析で新陳代謝度は回帰係数、p-value、R2 いずれの統計量においても低い評価であったため、説明変数から除外することとした。はじめに多重共線性を回避するため、VIF による説明変数の選択を行ったところ、P 値、次数中心性 (Degree)、ページランク (Pagerank) を削除すれば全ての説明変数において VIF が基準値の 5 以下になることがわかった (表 5.14)。

表 5.14: VIF と説明変数の選択

	Z	P	PW	Degree	Closeness	Between	Pagerank	Burt	Neighbor
VIF1	1.071	11.917	5.817	189.287	17.346	7.581	174.153	4.591	12.676
VIF2	1.03	-	1.577	-	3.540	1.382	-	2.845	4.986

以降の分析ではこの VIF の基準に基づいて選択された 6 つの指標、Z 値、近接中心性 (Closeness)、Burt の拘束指標 (Burt)、取引数 (Neighbor)、媒介中心性 (Between)、および PW 指標を用いて重回帰分析を行う。各説明変数の組み合わせによるモデル選択は、VIF、p-value および回帰係数を基準として 7 通りのモデルを選択し、統計的検証を行った (表 5.15)。分析の結果、媒介中心性と PW 指標のみのモデルにおいても p-value が 0.01 以下の有意性を示し、PW 指標単独でもある程度の説明力はあることがわかった。すなわち、PW 値が高いほどクラスター継承度が低い傾向があることが確認された。

表 5.15: クラスタ継承度とノード評価指標の重回帰分析結果

	Model1	Model2	Model3	Model4	Model5	Model6	Model7
Z	-0.003						
Closeness	-0.223*	-0.223*					
Burt	0.232**	0.231**	0.232*				
Neighbor	0.545***	0.545***	0.372***	0.186**			
Between	-0.344***	-0.344***	-0.372***	-0.325***	-0.296***		-0.050
PW	-0.629***	-0.630***	-0.706***	-0.729***	-0.732***	-0.633***	
<i>Adj.R</i> ²	0.517	0.522	0.512	0.497	0.468	0.395	-0.007

p* < .10 *p* < .05 ****p* < .01

5.4.3 機械学習モデルによるクラスタ継承度の予測

本節では、比較的高い精度を出す回帰モデルとして汎用的に使用されている Elastic Net での正則化を適用した一般化線形回帰である GLM-net, Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost) を用いて各説明変数の重要度を比較する。まず、それぞれのアルゴリズムにおいて、ハイパーパラメーターのグリッドサーチを行い、5-fold の交差検証でモデル計算を行った。RMSE, R2, MAE の各統計量を比較した結果、全ての評価値において最上位モデルはランダムフォレストであった（表 5.16）。

表 5.16: 各アルゴリズムの統計量

Algorithm	RMSE	R2	MAE
GLM	0.636	0.485	0.467
RF	0.564	0.595	0.405
SVM	0.885	0.505	0.688
XGBoost	0.936	0.434	0.669

最上位モデルであったランダムフォレストにおける説明変数の重要度を比較した結果を図 5.21 に示す。各説明変数の重要度は最も重要な説明変数の重要度を 1 とした時の相対的な値として表示している。クラスタ継承度の予測において、重要度が最も高い説明変数は提案指標 PW であった。

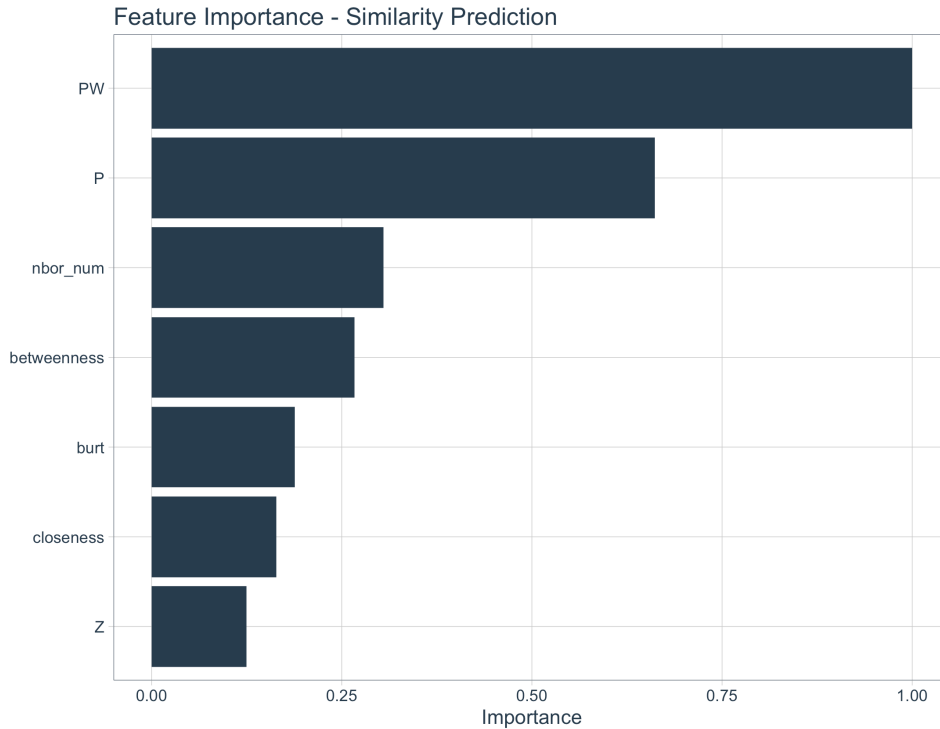


図 5.21: クラスター継承度に対する各説明変数の重要度

5.5 考察と結論

外部の知識を吸収して適応する企業は、特に予期せぬ環境変化に対してより強いレジリエンスを持つと考えられてきた。本章では、異質性の高い産業クラスターをつなぐことでネットワークに多様性をもたらす企業の仲介者の役割を探るため、東北地方の産業クラスターの進化を調査した。そのため、提案手法に基づいてブローカーの役割を果たす企業を抽出し、抽出された企業のネットワーク的な特徴を調べると共に、10年間の間にブローカー企業がどのようなクラスターに所属するかを調査した。企業間取引ネットワークの動態を定量的に分析した結果、ブローカー企業は隣人が少なく、遷移が多く、クラスター間の移動性が高い傾向にあることがわかった。また、ネットワーク構造を反映する指標別にクラスター進化マップを作成し、各クラスターパスの特徴を抽出した上で、そのパス上に存在する代表的な企業の役割の違いを示した。

クラスター進化マップの観点から、3つの特徴的なクラスターパスが浮かび上がってきた。第一のクラスターパスは、東北の全県をカバーする地域的な多様性を持つ内部リンクが連続的に存在する安定的なクラスターである。このクラスターパスの企業は、主に水産業を中心とした企業であり、取引先の入れ替わりはほとんどなく、希少な情報をもたらすブローカーがほぼ見られなかった。第二のクラスターパスは、ローカルハブを持つ地域クラスターである。このクラスターパスの安定性は、建設業のビジネスリンクが多い巨大なハブ企業によって維持されていることが示唆された。一方で、このクラスターパスでは、多くのブローカー

を持つ小さなクラスターが、より大きなクラスターに合併したり、クラスターが分裂するといった進化イベントが多数検出されており、ハブ企業との関係が薄いクラスター周辺部における安定性は低いと言える。第三のパスは、多くのブローカーを持つ小規模なアジャイルクラスター群から構成されていた。このクラスターパスでは2009年以降は連続的に継承クラスターが存在し、大きな進化イベントは検出されなかったが、新陳代謝度、PW値共に他のクラスターパスに比べて非常に大きな値であった。このクラスターパスの主要地域は、東日本大震災の被害を大きく受けた後、復興が進み、現在高い経済活力を持つ宮城県であることを踏まえると、この地域においてネットワークのリワイヤリングが効果的になされたことが示唆される。

上記の3つの事例の観察は、本章で検出されたブローカーの特徴である「隣人が少ない」「移行性が高い」「移動性が高い」という点で一致している。しかし、こうしたブローカーの特徴は、ネットワークのダイナミクスの中でのブローカーの活動を捉えることを困難にしている。ブローカーが多いクラスターほど進化的事象の経験が多いように見えたが、クラスターの進化におけるブローカーの役割や外部知識の吸収を検証するものではない。今後はネットワークリンク上の情報をフローとして取り込むランダムウォークベースのモジュール分解[192]など、より細かいスキームでの分析を取り入れていく予定である。

第6章 考察

企業の新陳代謝や離れたコミュニティに存在する異質な情報へのアクセスは、理論研究やケーススタディにおいてはイノベーションに重要な役割を果たすことが認識されている。本論文では取引関係の入替率とコミュニティ構造の異質性というかたちでこれらの概念を指標化し、企業の存続年数やクラスター進化に関する企業のデータ解析に活用した。本節では提案指標を用いた解析結果を振り返り、本論文のアプローチから導き出される示唆や、政策現場への適用可能性について議論する。

6.1 変動する市場環境への適応力に関する考察

6.1.1 企業取引の新陳代謝度の企業パフォーマンスへの影響

IT バブル崩壊、リーマンショック、東日本大震災に伴うサプライチェーン崩壊といった予測できない大規模な経済ショックが近年頻繁に発生している。変動が激しく、変化スピードが速い近年の市場環境は、企業の環境適応力を測定することの重要性が高まっていることを示唆している。県別 GDP や業種別売上高といったこれまでの静的な地域経済指標とは異なる、新しい動的な定量指標が必要である。第3章では、東北地方、中部地方、九州地方の大規模な企業間取引データを用いて、企業の新陳代謝を定量的に測定することに初めて成功した。解析の結果、取引が固定的な企業は存続年数が短く、市場から撤退する傾向があることが判明した。新陳代謝度と企業の存続年数の統計的解析結果は、取引先企業を毎年1~2割程度変更する企業群の存続年数が最も長く、取引先企業をほぼ変えないか大幅に変える企業、すなわち新陳代謝度が極端に小さいか極端に大きい企業の存続年数はいずれも短くなる傾向があることを示していた。取引ネットワークの新陳代謝に関する本論文の客観的な解析結果は、地域的なイノベーション環境や企業の経営革新の程度を把握するための有用な情報となりうる。新陳代謝度は企業間取引のダイナミクスを捉え、地域経済の動向を知る新たなマクロ経済指標としての利用が考えられる。

6.1.2 拮抗する二つの力

イノベーションに成功するネットワークは、拮抗する二つの力のバランスを取っていることが知られている [114]。一方では、ネットワーク構造が柔軟で新規参入者や情報の吸収力が高いことが重要とされ、企業間関係が固定的でクラスターの境界が変化しないネットワークはイノベーションを生みにくいとされる。しかしもう一方では、ネットワーク構造があまり

に不安定だと、企業間に信頼関係が生じにくく、技術革新に関する重要なアイデアが共有されないため、企業が新たなコンピテンシーを開発することが困難になる。村上が生物進化と遺伝子の複製機能の関係を参照しながら経済成長の動学的モデルを考察して、「普遍性を貫こうとする力とそれを変える力が絡み合う二重の機制が進化である」と述べたように [137]、ネットワークの進化を分析する際にも、ミクロレベルで拮抗する二つの力を考慮することは重要である。本論文の新陳代謝度に即して述べるならば、企業レベルでも既存の取引先を維持しようとする力と取引先を変えようとする力という二つの力が働いていると言える。存続年数の長い企業の新陳代謝度が1割程度であったということは、取引を維持しようとする力の方が変えようとする力より強いことを示しており、二つの力のバランスは維持する方に傾いている。このことは、変わらないものが変わっていくという逆説的なバランスが、生物界だけでなく、企業間ネットワークの進化の原動力にもなっていることを示唆している。

6.2 異質性を取り入れる仕組みとイノベーションの可能性

6.2.1 ブローカーの検出手法

地域クラスター内部の企業が企業間の密なコミュニケーションによって専門性を高め、クラスターに属さない企業に比べて競争上優位である状態が行き過ぎると、イノベーションを妨げる負のロックインに陥ってしまうことが知られている [65]。新しいイノベーションの機会は多様な技術や知識源に触れることでもたらされることが多いが、そのためにはクラスター間の移動や交流が促進され、クラスター外の異質な情報に対して開かれた状態であることが望ましい [66]。第4章ではクラスター差異を反映したノード指標PWを提案し、ネットワークの構造的空隙を埋め、異質なクラスターを結びつけるブローカーを検出する手法の有効性を検証した。コミュニティ間の距離が異なる3種類のネットワークにおけるノード順位の可視化により、提案指標PWが遠く離れたコミュニティをつなぐノードを検出する指標として、比較に用いた6種類のノード指標の中で最も優れていることがわかった。また、東北ネットワークを用いて、ノード削除後の平均最短経路長を比較したところ、ノード削除によって失われるエッジ数に対して、最も大きな影響力を持つノードを検出するのは、7つの比較指標よりPW指標の方が優れていることがわかった。すなわち、PW指標は稀少かつ重要なリンクを持つノードを検出するのに最も優れたネットワーク指標と言える。

6.2.2 ノードのポジション価値とクラスター進化

ネットワーク中のノードのポジション価値を重視する分析では、ハブのような多数のエッジを持つノードが注目される場合が多いが、数は少なくとも重要なエッジを持つノードを見逃さないことも大切である。そのようなノードはクラスターの周辺に位置し、クラスターに希少な外部知識を運んでくるニッチプレイヤーである可能性が高いからである。ニッチプレイヤーは異質なクラスターを結びつけるとともに、容易に入れ替え可能な緩やかな連結により、ビジネスエコシステムの変動への柔軟な対応を可能にする [57]。第5章では、PW指標

を東北地方の企業間取引ネットワークに適用し、ブローカー企業を検出してその特徴を分析した。その結果、ブローカー企業はハブ企業のように取引先を多く持たず、移り変わりとクラスター間の移動性が高いことがわかった。提案指標 PW はクラスターにおいてニッチプレイヤーの役割をもつ企業を検出していることが示唆される。

次に着目したのは、ネットワークのミクロレベルにおけるノードの役割や動向が、マクロなクラスター間の関係やクラスター進化に及ぼす影響である。クラスターの進化イベントを識別するこれまでの分析手法では、ネットワークの中でどのようなポジション価値を有するノードが、クラスター進化に寄与するのかが不明であった。第5章では、ネットワーク中のノードのポジション価値を評価する9つの指標を用いて、クラスターの進化イベントを識別する際の基準値である、クラスター継承度の予測を行った。その結果、クラスター継承度の予測に対する寄与度はPW指標が最も大きかった。この結果を踏まえると、PW指標が検出するブローカー企業は、地域の産業構造を進化させていく触媒となる重要な役割を果たしている可能性が高いと言える。

6.3 提案手法の組み合わせによる新たな示唆

本論文における2つの提案指標である新陳代謝度とPW指標を用いて行ったクラスター進化の分析により、企業のネットワークにおけるポジション価値は、その企業が属するクラスターパスによって大きく異なることがわかった。本節では第5章で論じたビジネスエコシステムの多様性を維持する仕組みについて考察し、本論文の分析結果から考えうる政策的インプリケーションを述べる。

6.3.1 越境型クラスターと多様性維持の仕組み

第5章では、クラスター進化マップとクラスターパスのレーダーチャートを用いて、東北地方におけるクラスターの動的推移過程やネットワークの局所的構造といった多面的な特徴を把握する方法を示した。これらは以下の点で、従来のクラスター研究とは異なる視点を提示するものである。

Porterのクラスター戦略は、集積効果による産学連携の促進など、クラスター一般に見られる利点に着目して展開されており、クラスターを特定の産業に結び付けて捉えるアプローチであった[32]。これは政策資源の戦略的な投資対象を定める際の根拠として用いられやすく、多くの先進国がいち早く取り入れたアプローチでもある[37]。しかし、異質な分野の融合が進み、従来の産業区分が企業活動の実態から離れてきている現代の市場構造を考えると、単一産業で構成されるクラスターの競争優位に注目する観点は、実社会への適用面で限界がある。第5章の分析により、東北ネットワークのクラスターは産業エントロピーや地域エントロピーが互いに大きく異なり、単一の産業ないし地域に属する企業が大半を占めるクラスターが存在する一方で、多産業・多地域の企業で構成される越境型クラスターも多く存在することがわかった。

このようなクラスター自身の多様化や環境適応に加え、クラスター外の企業やクラスターとの相互関係、およびそれらの変容に伴うクラスター進化イベントも、クラスターの特性を捉える上では重要である。第5章の分析は、クラスターをノード指標に基づいて様態別に捉えている点、クラスターの年次推移に着目したクラスターパスの分析でそのダイナミクスを捉えている点で、従来のクラスター解析で見過ごされてきた重要な観点を補うものと言える。

6.3.2 政策的インプリケーション

データとデータサイエンスの手法を用いた政策の立案や施策の運用

政策現場におけるデジタルトランスフォーメーション（Digital Transformation：DX）の先駆的な例としては、内閣府エビデンスシステム（Evidence data platform constructed by Council for Science, Technology and Innovation：e-CSTI）、内閣府地域経済分析システム（Region economy society analyzing system：RESAS）、地域未来牽引企業の選定における帝国データバンクのデータやコネクター・ハブ指標の活用等があげられる。しかし、こうした一部の事例を除き、現状では、社会におけるDXのスピードに政府が追いついていない。先駆的な例を社会実験として、データ活用を拡げていくべきである。本論文を通じて明らかになった、政策立案・評価の観点からの具体的な知見・メリットとして以下が挙げられる。

- 地域に形成された企業社会は絶えず変化しており、データを用いたアプローチなしに、そのダイナミクスを把握することは困難である。
- 地域には、隠れた重要企業群が多数存在している。従来型の人による訪問のアプローチでは、時間的、コスト的に、それらを見出すことは困難だが、データを活用することで、候補を絞り込むことが可能である。
- 地域に形成された企業集団やクラスターパスは多様である。多様性を理解した上で施策投入を行うことで費用や時間対効果を上げうる。
- 時系列でクラスターパスや重要な企業の動きを追っていくことで、政策評価に有益な情報が得られる。

生態系としての企業間ネットワークと地域産業経済圏

企業は地域に単独で立地しているわけではなく、相互に結びついたビジネスエコシステムとしての生態系の中に生息している存在である。しかし、企業の活動や成長に大きな影響を与える産業の生態系の姿は、例えば、地域別の事業所数、出荷額、地域の産業特化係数といった、従来用いられてきたデータや指標で十分に捉えることは難しい。地域経済や社会の実態を産業の生態系として捉えるには、ネットワーク分析の手法を用いることが有効である。

ネットワークの考え方が施策に取り入れられた初期の代表例として、東日本大震災の復興対策の一つである「グループ補助金」がある。これは、サプライチェーンを構成する企業のなかで、グループ単位で協力して工場設備や旅館の復旧をしようとする集団に対し、優先的に復旧支援を行う施策である。サプライチェーン全体として復旧に取り組むことにより、設備投資を行ったが部品や販売先が無いために、商品を生産・出荷できないという事態を防止することができ、被災地では、高い評価を受けた施策となっている。

本論文では、地域の産業構造を表す新しい手法として、企業間に形成された取引ネットワークを客観的に捉え、既存の産業分類や地域の境界を超えて有効に機能する企業群の抽出と、それらの比較分析を可能にする指標や枠組みを提案した。以下では政策形成支援の観点から、提案手法の具体的活用方策について論ずる。

(1) クラスターパスの特性に応じた施策の投入

本論文を通じて、地域に形成された企業集団やクラスターパスは、それぞれ異なる特徴を持つことが明らかとなった。ここでは前節で論じた、多様性を理解した上で施策投入を行うことで費用や時間対効果を上げることの具体例として、本論文から得られる政策的インプリケーションについて述べる。

クラスターパス 1,2,3 はいずれも 10 年ないし 8 年間の長期にわたり、クラスターペアの類似度が高い状態が続いた比較的安定なクラスターパスである。クラスターパス内の各継承クラスターは期間を通じて同様の特性を示したが、クラスターパス間の特性は大きく異なっていた。ここでの特性とは、第 5 章で示したレーダーチャートの 8 つの特性、すなわち、新陳代謝度、Z 値、P 値、PW 値、売上高、産業エントロピー、地域エントロピー、クラスターサイズのことである。この結果を踏まえると、クラスターパスの特性に応じた政策的支援策を昂じることが有効であることが示唆される。以下では 3 つのクラスターパスについて、それぞれ考えうる政策的インプリケーションを述べる。

1. 地域多様性を持つ定常クラスター

PW 値が非常に低いこのタイプのクラスターでは、域内取引で企業間関係が完結している場合が多く、新陳代謝度も極端に低いため取引が固定化しやすい。安定してはいるものの、新たなイノベーションを伴う将来的な発展が見込みにくいと言える。異質な外部知識を取り入れ多様な技術資源を活用できる制度環境を有する企業に比べ、企業間関係のダイナミクスに乏しく特定技術への依存性が高い企業は、負のロックインに陥ることが知られている [65, 193]。このような場合には、域外取引のポテンシャルを引き出しうる政策が必要である。その際、クラスター内企業の中からクラスター外の知識や技術に対する吸収力が高い企業を見分け、有力なニッチプレイヤーとなれるよう支援することが有効であると考えられる。ニッチプレイヤーへの投資はクラスターに希少な情報や技術、産物をもたらすとともに、大市場へのパスをつくる契機ともなりうるだろう。

2. ローカルハブを中心とした地域クラスター

一つの地域に特化して存在するこのような地域クラスターでは、地域経済の命運を決めるネットワークのハブである城主企業の動向をモニタリングする必要がある。地域未来を牽引する企業に政策資源を集中させるアプローチは、従来日本の産業政策でも行われてきたが、城主企業が経済ショックに対して脆弱な場合は、クラスターが大きくその影響を受けて縮退する可能性もある。合併や分裂といったクラスター進化のイベ

ントも多く見られることから、城主企業から離れたクラスターの周辺部において、企業の出入りが激しく、知識の流動性が高いことが推察される。このようなネットワークでは、相互に連結した複数のハブ企業で構成されるネットワークの中心部と、主に2社関係で構成される周辺部が区別されるという特徴がある。そこでは、従来の技術と新しい技術を組み合わせ、既存の製品を変革する生産システムをどのように構築するかが課題となる [194]。また、クラスターを構成する業種は特定産業に偏る傾向があるため、他分野進出の可能性を追求しつつ、将来的なショックを緩和しうるネットワーク形成への支援が有効だと考えられる。具体的には、海外市場紹介や産学協働での研究開発、国際見本市など、知的対流の場を生み出す仕掛けが必要となるだろう [195]。

3. ブローカーが局在する小規模アジャイルクラスター

新陳代謝度が非常に高いこのタイプのクラスターは、市場変動に柔軟に対応する能力を持っているものの、安定的な成長力に欠けるため人材育成や組織学習、経営基盤構築のための支援が有効になると考えられる。ブローカーが多く存在するクラスターであるが、ブローカーは移り変わりと移動性も高いため、個別企業をターゲットにした政策支援は困難である。クラスター外から知識を獲得するブローカーは、知識の創造や拡散に積極的に貢献するゲートキーパーの役割を果たす可能性が高い [196, 74]。一方で、こうした異質な知識ベースを持つ企業は、ローカルコミュニティへの参加を通してパフォーマンスを向上させることが知られている [197]。このような場合は、クラスター外の知識や技術をクラスター内部に根付かせる知識係留の強化や、ネットワーク形成を促し成長軌道に載せていくための面的支援が必要になる。産学官金融との連携促進や、ローカルメディア等を通じた魅力発信、多様なリソースを結びつける異業種マッチングへのインセンティブ付与も重要な施策になると考えられる。

(2) データを用いたアプローチによる「ブローカー企業群」の特定

第5章では、クラスター進化に関して、「ブローカー企業群」が果たす役割の重要性を指摘した。クラスターを市場環境の変動に呼応してアップグレード [31, 32] するためには、クラスター内の企業の自助努力に加え、外部性をもたらすブローカー企業群の役割が重要となる。しかし、人の訪問や面談によるアプローチでは、地域に多数存在するブローカー企業群を把握することは困難である。ブローカー企業群は、地域のまとめ役として重要なハブ企業とは異なり、取引関係の変動に応じて該当企業が変化する傾向が顕著であり、仮に一度、把握できたとしても年単位で更新を行うことが必要となる。本論文が提案したPW指標により、ブローカー企業群を抽出し、分析することが有効であることが見込まれる。

(3) 安定的な企業群と変化を特徴とする企業群とのマッチング

第3章における企業の新陳代謝についての分析結果を踏まえ、6.1節では、企業間ネットワークの進化には、取引関係を維持しようとする力と変えようとする力という、二つの拮抗

する力が働くことが重要であることを論じた。これに対し、第4章、第5章では、ネットワークの構造的特徴に基づいて抽出したブローカー企業が変化しやすいこと、クラスター進化に最も大きく寄与することを示した。また、ブローカー企業が安定的に存在するクラスターでは、ハブ企業とブローカー企業のリンクが多いことを示した。これは、安定的なハブ企業と変化を特徴とするブローカー企業が共に働くという点で、二つの拮抗する力が、ネットワークの進化に重要な役割を果たしていることのもう一つの側面と言える。

すなわち、産業の生態系では、動的推移、静的構造特徴の双方において、安定と変化という二つの力のバランスが重要であることが明らかとなった。一方、現在、地域に存在する複数の企業から構成される組織である商工会議所や商工会は、安定的な企業群の間の交流に偏っていることが知られている。クラスター進化を誘発するためには、安定的な企業群と変化を特徴とする企業群との出会いの機会を増やすことが有効であると考えられる。具体的な方策として、既に地域市場で把握されているハブ企業をはじめとする安定企業群と、データを用いて探索したブローカー企業群とのマッチングの機会を設定することが考えられる。

6.4 手法の限界と今後の研究課題

6.4.1 データおよび指標設計上の限界

本論文で用いた企業間取引データは、帝国データバンクの場合は各企業につき発注社・受注社それぞれ5社まで、東京商工リサーチの場合は各20社までの主な取引先企業が集められたものであり、各企業の全ての取引先企業を網羅しているわけではない。データは毎年更新されているが、その年において各企業が主要な取引先だと考える企業があげられており、主要な取引先とされない企業はデータに含まれないため、分析対象外となっている。また、本論文の分析対象企業は、取引先10社以上もしくは取引先8社以上の企業に限定しており、取引先企業数が少ない企業は分析対象外となっている（新陳代謝度と存続年数の相関関係については、東北地方の全企業を対象にした分析結果を補足資料A.1に掲載している）。

提案指標である新陳代謝度は、各企業における取引先入替率の期間を通じての平均値として定義しているため、時系列の変動を扱うことはできない。また、定義上、取引先企業が増えた場合、入れ替わった場合、減った場合の全ての場合において新陳代謝度が増加するため、新陳代謝度が変動した時の解釈においてこれら3つの要因を区別することはできない。この課題を乗り越えるため、第3.3節では、取引先企業数と被消失取引数の特徴量に加えた非負値テンソル因子分解を行った。この手法は各特徴量の分位点水準における企業数に基づく特徴抽出を行うものであり、データの中に潜在するパターンを抽出して解釈することは可能だが、新陳代謝度の変動要因を説明することはできない。

提案指標PWに用いたエッジの重み付けは、情報エントロピーの考え方を参照し、コミュニティ間距離に反映されるリンク確率の平均値を考慮して設計した。その際、コミュニティ間距離はコミュニティ間のエッジ数と共通隣接ノード数に基づいて計算しているが、他にもコミュニティ間の関係を表すさまざまな距離の計測方法が考えられる。今後はランダムウォークに基づくノードのコミュニティ滞在確率の推定等 [192, 198]、より精緻なコミュニ

ティ間関係の分析が検討課題となる。また、本論文では、売上高や利益率、従業員数といった企業固有の特微量は限定的な利用をしており、ネットワークの構造情報にノード自身が保有する特微量をどのように結びつけるかは今後の研究課題である。

第5章で行ったコミュニティ進化の動態分析では、年次毎のネットワークのスナップショットの集合をマッチングさせることで、ネットワークの変動を捉えている。その際、コミュニティ分割は所与のものとし、始めに行ったクラスタリング結果をそのまま適用させている。しかし、本論文で採用したクラスタリング手法である Louvain 法は、モジュラリティ最大化に基づく貪欲型のアルゴリズムであり、最適解（大域最大）を保証するものではない [199]。Louvain 法は高速で精度も高いため、広く用いられているが、モジュラリティ最大化に基づく他手法と同様、クラスタリング結果は実行毎に微妙に異なるため、留意が必要である。実際、第5章の解析において、別途実行したクラスタリング結果を適用させたところ、非常に安定的なクラスターパス 1,2 では同様の特徴が得られたが、クラスターパス 3 は異なる特徴を示した。クラスタリング精度の向上は本論文の目的の範囲外だが、クラスター間関係の微細な変動を分析するためには外せない観点であり、将来的な研究課題としたい。

6.4.2 提案手法の適用対象と影響伝播モデルへの応用

コミュニティ間距離を定量化してノード評価に活用する提案指標 PW は、多様なコミュニティ構造を持つさまざまなネットワークに適用することが可能である。たとえば、論文の引用ネットワークに適用して、異分野融合の分析等に適用することも可能であるし、ソーシャルネットワークに適用し、近年問題が指摘されているエコーチェンバーやフィルターバブルに見られるような、同質の情報が増強・強化されてしまう現象への戦略的対応策として、ネットワークに異質性をもたらす媒介者に着目する場合等に利用することも考えられる。

提案指標 PW における稀少かつ重要なエッジの評価は、情報伝播・拡散モデルにおいて、ノードに以下のような活性化や不活化の条件を課す場合に有用であることが見込まれる。一つは、社会ネットワークにおける情報カスケードである。各ノードはある情報量を受け取ると活性化し、その情報を伝播するとする。次数が高いほど隣接ノードから受け取る情報の稀少価値は下がるため、ハブよりは PW 値の高いノードの方が活性化されやすく、大域的なカスケードを引き起こしやすい可能性がある。もう一つは組織的障害からの回復などの異常時対応である。問い合わせが殺到するなど情報過多になるとノードが不活化し、情報伝播力が弱まるとする。このような場合は、PW 値の高いノードに情報を提供する方が不活化による損失が広がりやすく、ネットワーク全体に情報が共有されやすい可能性がある。情報拡散モデルは感染モデルにも類似しているが [200]、ネットワークのコミュニティ構造やノードのポジション価値を、情報拡散モデルや感染モデルに組み込んだアプローチはまだあまり知られていない [201, 202]。今後はこうした複雑ネットワークの構造に基づく影響伝播に着目した分析への応用も考えられる。

第7章 結論と展望

7.1 本論文の結論

市場環境の変動に対する企業の適応力を定量的に評価する指標の不足は重要な課題である。本論文の目的は、地域経済政策やイノベーション研究の中で理念的に論じられることの多かった「新陳代謝」「ニッチ」「クラスター進化」といった概念を検証可能な形で新たに定義し、ネットワーク科学の手法を使ってその妥当性を検証するとともに、経済活動のダイナミクスを定量的に解析することであった。この目的に対する本論文の結論について述べる。

本論文では、企業間取引ネットワークにおける企業のミクロな取引先交代を計測する新たな指標として新陳代謝度を開発し、日本企業の新陳代謝度を定量的に測定することに初めて成功した。解析の結果、大半の企業が取引先企業を毎年1割程度変更していたことが明らかになった。また、取引が固定的な企業および取引関係の変動が過度に大きい企業は、市場での存続年数が短い傾向があることを示した。次に、不均質なコミュニティ構造を持つネットワークにおける、エッジの稀少性を評価するノード指標PWを開発した。条件を変えて生成した複数のベンチマークネットワークを用いて、ネットワークの攻撃耐性に関する検証実験を行った結果、提案指標PWは比較指標よりも、稀少性の高いエッジを持つノードを識別する上で優れていることが示された。提案指標PWを東北地方の企業間取引ネットワークに適用したところ、PW値が高い企業はビジネスエコシステムの変動への柔軟な対応を可能にする、ニッチプレイヤーとしての特徴を有していることがわかった。また、クラスター進化イベントへの寄与度は提案指標PWが最も大きいことが示された。最後にケーススタディとして、典型的な3つのクラスターパスを分析し、2つの提案指標を含む8つの指標とクラスター進化イベントの出現傾向において、クラスターパス毎に異なる特徴があることを明らかにし、その特徴を踏まえた政策的支援方策について論じた。

以上から、本論文はネットワークを構成する企業の動的推移と構造特性を捉える2つのノード指標を開発し、その有用性を示すとともに、企業が所属するクラスター進化のダイナミクスを、提案手法に基づいて分析する有効な方法の開発に成功した。

7.2 今後の展望

今後の展望としては、本論文で開発したノード指標およびクラスター進化の分析手法を用いて、広域経済圏におけるクラスター間関係を把握し、特徴の異なる複数のクラスターパスを組み合わせた、典型的なポートフォリオを描くという応用が考えられる。変動の激しい市場環境において経済圏の持続可能な発展をなすためには、安定的なクラスターばかりではな

く、柔軟性の高いアジャイルクラスターや成長力のある小規模クラスターなど、異なるクラスターパスの動的特性を踏まえた、多様なクラスターの組み合わせを考慮する必要がある。生産要素の移動性や知識係留を考慮した多地点ネットワークのダイナミクス解析においても、ハブやブローカーの局在といったクラスターの様態に基づくメゾレベルの種別、さらには複数のクラスターパスのバランスといったマクロレベルの動態解析は有効であることが推察される。

本論文で直接的には取り上げなかったが、情報通信産業と科学研究・専門技術サービス業は、どちらもクラスター間の移動性が高く（クラスターパスを外れることが多く）、かつ、同じクラスター内に位置していることが多かった。近年の情報通信産業における技術革新スピードの上昇を考えると、科学研究・専門技術サービス業と連携し、その成果を積極的に取り入れることが要因の一つになっている可能性がある。今後は大学と企業との共同研究を含めた柔軟な連携関係のネットワーク分析の充実が望まれる。現状では、大学側に情報守秘の意識が高まったことで解析可能なデータに制約が生じており、産学連携の共同研究に関する分析は進んでいない。しかし、中央研究所等に力点を置いた単独の研究体制から、大学や他の研究機関との連携体制への移行を模索する企業が増える中、共同研究を定量的に分析し、客観的に評価することの必要性は高まっている。

企業間あるいはクラスター間における、競合関係の分析も重要なテーマである。愛知県の新陳代謝度が経済規模に比べて小さいのは、城主企業であるトヨタを中心としたサプライチェーンの頑健性を表すものであり [203]、取引関係が途切れないことによって守られている中小企業が多数存在していることを示唆している。しかし、このような構造が長期的視野で見た時の持続的な地域経済の発展にとって有益かどうかは不明である。本論文では割愛したが、本来は競合関係にあると思われる、取引先企業が類似している企業群において、新陳代謝度が著しく低いのは協同組合であった。日本企業に見られるこうした互助関係や取引維持の仕組みについては、データに基づいた分析を行うと共に、実際のフィールドワークを通してその実態や背景にある要因を探る必要がある。本論文で開発した手法は、そのようなフィールド調査を実施する際の、重点対象の絞り込みや質問項目等の設計においても役立つことが期待される。

謝辞

本研究を進めるにあたり、指導教員である東京大学大学院工学系研究科技術経営戦略学専攻の坂田一郎教授には、貴重な時間を割いて親身にご指導いただきました。私に研究再開の道を開いてくださり、新たな分野で学びながら研究する稀有な機会と環境を与えてくださったことに心より感謝いたします。また、研究の方向性や意義付け、研究成果のまとめ方にいたるまでの丁寧なご指導、坂田先生の幅広い知見と経験に基づくご助言は、本研究の大きな糧となりました。深く感謝いたします。

本論文の審査をお引き受けくださいました、東京大学大学院工学系研究科システム創成学専攻の和泉潔教授、東京大学大学院工学系研究科技術経営戦略学専攻の田中謙司准教授、東京大学大学院情報理工学系研究科附属情報理工学教育研究センターの森純一郎准教授、東京大学公共政策大学院の芳川恒志特任教授には、研究の意義や手法の限界、分析結果の評価等に関する重要なお指摘をいただきました。あらためてお礼申し上げます。

東京大学イノベーション政策研究センターの浅谷公威特任助教、大知正直特任研究員、東京大学未来ビジョン研究センターの佐々木一特任准教授には、共同研究や研究会での議論を通して、多くのご指導をいただきました。磯沼大さん、三浦崇寛さんはじめ、坂田・森研究室の皆様からは研究会での議論等を通して大変刺激を受けるとともに、本研究の分析結果に関する重要なお助言をいただき、研究内容を深めることができました。本研究室に在籍されていた郷治友孝さん、古瀬利博さんからは、共に研究を進めるなかで励みになる言葉を度々かけていただきました。誠にありがとうございました。

東京大学大学院工学系研究科システム創成学専攻の鳥海不二夫教授、株式会社 ALBERT の松林達史先生には、ネットワーク生態学研究会での議論をはじめ、研究内容に関する重要なお助言をいただきました。特に松林先生には、提案指標の開発の初期段階からアドバイスをいただき、大変力付けていただきました。京都大学こころの未来研究センターの広井良典教授からは、本研究の着想のヒントとなった持続的社会モデルへの試案について、大変示唆に富んだご助言をいただきました。広井先生の本研究へのご関心と温かいメッセージは、研究を進める上で大きな励みになりました。東京大学大学院工学系研究科バイオエンジニアリング専攻の岡本洋特任研究員には、浅谷特任助教と共同で立ち上げた論文勉強会を通じて、ネットワーク科学分野の最新の研究動向について貴重な知見をご提供いただきました。心から感謝いたします。

本研究を進める上で、数多くのインスピレーションと知的刺激を与えてくださった、東京大学大学院理学系研究科の五神真前総長、東京大学大学院人文社会系研究科の佐藤健二副学長はじめ、総長ビジョン検討委員会のメンバーでいらした先生方に、深く感謝いたします。本学の学生達のために、多様な専門分野の知のプロフェッショナルの先生方が、多くの議論

を重ねて紡ぎ出される言葉は、私自身の研究生活の指針になりました。東京大学未来ビジョン研究センターの城山英明センター長はじめ、同センターの皆様のお力添えで、研究者としての道に再び進めるようにしていただきましたことを、誠に感謝いたします。東京大学未来ビジョン研究センターの前身である、東京大学政策ビジョン研究センターの設立時のメンバーであり、広報担当の職員だった私に、研究再開のきっかけとなる共同研究のお誘いを最初にくださった、埼玉医科大学の中島勸教授に、心から感謝申し上げます。

本研究室の秘書の石原絢さんと佐藤妙子さんには研究に専念できるように多くの事務手続きを行なっていただきました。東京大学未来ビジョン研究センターの三浦則男さん、小林範子さん、村上壽枝さん、中村裕子さんをはじめ、事務局の皆様にはさまざまな形でご支援いただきました。皆様のお力添えに心から感謝いたします。

最後になりましたが、研究生活を支えてくれた家族と友人、関係者の方々に心より深く感謝いたします。幼い頃から研究の楽しさに触れさせ、見守ってくれた亡き父に、感謝の報告をしたいと思います。本当にありがとうございました。

参考文献

- [1] Shin-Ichi Yamamoto. Forest gap dynamics and tree regeneration. *Journal of forest research*, Vol. 5, No. 4, pp. 223–229, 2000.
- [2] 阿部博史, NHK スペシャル「震災ビッグデータ」制作班. 震災ビッグデータ: 可視化された〈3・11の真実〉〈復興の鍵〉〈次世代防災〉. NHK 出版, 2014.
- [3] David J Teece, Gary Pisano, and Amy Shuen. Dynamic capabilities and strategic management. *Strategic Management Journal*, Vol. 18, No. 7, pp. 509–533, 1997.
- [4] Yahaya Y Yusuf, Angappa Gunasekaran, Ahmed Musa, Mohammed Dauda, Nagham M El-Berishy, and Shuang Cang. A relational study of supply chain agility, competitiveness and business performance in the oil and gas industry. *International Journal of Production Economics*, Vol. 147, pp. 531–543, 2014.
- [5] Wesley M Cohen and Daniel A Levinthal. Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, pp. 128–152, 1990.
- [6] Ben L Kedia and Rabi S Bhagat. Cultural constraints on transfer of technology across nations: Implications for research in international and comparative management. *Academy of Management Review*, Vol. 13, No. 4, pp. 559–571, 1988.
- [7] Y Sheffi and Rice JB Jr. A supply chain view of the resilient enterprise. *MIT Sloan management review*, Vol. 47, No. 1, pp. 41–48, 2005.
- [8] Belarmino Adenso-Diaz, Carlos Mena, Santiago Garca-Carbajal, Merrill Liechty. The impact of supply network characteristics on reliability. *Supply Chain Management: An International Journal*, Vol. 17, No. Issue 3, pp. 263–276, 2012.
- [9] Benjamin R Tukamuhabwa, Mark Stevenson, Jerry Busby, and Marta Zorzini. Supply chain resilience: definition, review and theoretical foundations for further study. *International Journal of Production Research*, Vol. 53, No. 18, pp. 5592–5623, 2015.
- [10] Nils-Ole Hohenstein, Edda Feisel, Evi Hartmann, and Larry Giunipero. Research on the phenomenon of supply chain resilience. *International Journal of Physical Distribution & Logistics Management*, Vol. 45, No. Issue 1/2, pp. 90–117, 2015.
- [11] 中小企業庁. 2006 年度版 中小企業白書, 2006.

- [12] Tetsuji Okazaki. Industrial development, firm dynamics and patterns of productivity growth: The case of the cotton spinning industry in prewar Japan, 1894–1924. *Research Institute of Economy, Trade and Industry (RIETI) Discussion Paper Series*, 2008.
- [13] 岡崎哲二. 経済史から考える: 発展と停滞の論理. 日本経済新聞社, 2017.
- [14] 中小企業庁. 2011 年度版 中小企業白書, 2011.
- [15] 中小企業庁. 2014 年度版 中小企業白書, 2014.
- [16] 総務省. 平成 26 年版 情報通信白書. 2014.
- [17] 中小企業庁. 2019 年版 中小企業白書. 2019.
- [18] 中小企業庁. 2016 年度版 中小企業白書, 2016.
- [19] 中小企業庁. 2020 年版 中小企業白書. 2020.
- [20] Charles H Fine. *Clockspeed: Winning industry control in the age of temporary advantage*. Perseus Books, 1998.
- [21] Clayton M Christensen. イノベーションのジレンマ 増補改訂版. 翔泳社, 2012.
- [22] Robert U Ayres. Industrial metabolism. *Technology and environment*, Vol. 1989, pp. 23–49, 1989.
- [23] Joseph A Schumpeter. *Theory of economic development*. Harvard University Press, 1911.
- [24] ヨーゼフ・シュンペーター, 塩野谷祐一, 東畑精一, 中山伊知郎. 経済発展の理論. 1977.
- [25] 鎌倉健. 産業集積の地域経済論-中小企業ネットワークと都市再生. 2002.
- [26] 山崎朗. 日本の産業クラスター計画: その背景と特質 (変化する日本の産業集積をめぐって). 経済地理学年報, Vol. 51, No. 5, pp. 499–511, 2005.
- [27] 経済産業省. 新産業創造戦略. 財) 経済産業調査会, 2004.
- [28] 松原宏. 日本のクラスター政策と地域イノベーション. 東京大学出版会, 2013.
- [29] Michael E Porter. The competitive advantage of nations. *Harvard business review*, Vol. 68, No. 2, pp. 73–93, 1990.
- [30] Michael E Porter. The economic performance of regions. *Regional studies*, Vol. 37, No. 6-7, pp. 549–578, 2003.
- [31] Michael E Porter. Clusters and the new economics of competition. *Harvard business review*, Vol. 76, No. 6, pp. 77–90, 1998.

- [32] Michael E Porter. *Clusters of innovation: Regional foundations of U.S. competitiveness*. Council on Competitiveness, Washington, DC, (Report.), 2001.
- [33] Mercedes Delgado, Michael E Porter, and Scott Stern. Clusters, convergence, and economic performance. *Research Policy*, Vol. 43, No. 10, pp. 1785–1799, 2014.
- [34] 後藤晃, 児玉俊洋. 日本のイノベーション・システム: 日本経済復活の基盤構築にむけて. 東京大学出版会, 2006.
- [35] 大久保敏弘, 岡崎哲二. 産業政策と産業集積: 「産業クラスター計画」 の評価. *Research Institute of Economy, Trade and Industry (RIETI) Discussion Paper Series 15-J-063*, 2015.
- [36] 松島克守, 坂田一郎, 濱本正明. 地域新生のデザイン. 東大総研, 2005.
- [37] OECD Publishing. *Innovative clusters: Drivers of national innovation systems*, 2001.
- [38] OECD Publishing. *The OECD innovation strategy: Getting a head start on tomorrow*, 2010.
- [39] Schumpeter. A rust-belt revival. *The Economist*, p. 56, 2016.
- [40] 西口敏宏. 遠距離交際と近所つきあい: 成功する組織ネットワーク戦略. NTT 出版, 2007.
- [41] 坂田一郎, 梶川裕矢. ネットワークを通して見る地域の経済構造—スモールワールドの発見 (特集ネットワーク最前線). *一橋ビジネスレビュー*, Vol. 57, No. 2, pp. 66–79, 2009.
- [42] 増田直紀, 今野紀雄. 複雑ネットワーク: 基礎から応用まで. 近代科学社, 2010.
- [43] Santo Fortunato. Community detection in graphs. *Physics reports*, Vol. 486, No. 3–5, pp. 75–174, 2010.
- [44] Mark S Granovetter. The strength of weak ties. In *Social networks*, pp. 347–367. Elsevier, 1977.
- [45] Ronald S Burt. Structural holes and good ideas. *American journal of sociology*, Vol. 110, No. 2, pp. 349–399, 2004.
- [46] Ronald S Burt. *Structural holes: The social structure of competition*. Harvard university press, 2009.
- [47] Roger Guimera and Luis A Nunes Amaral. Functional cartography of complex metabolic networks. *Nature*, Vol. 433, pp. 895–900, 2005.
- [48] Andrea Lancichinetti and Santo Fortunato. Community detection algorithms: a comparative analysis. *Physical review E*, Vol. 80, No. 5, p. 056117, 2009.

- [49] Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*, pp. 677–686. ACM, 2014.
- [50] Eli Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.
- [51] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pp. 36–43. ACM, 2005.
- [52] Kimitaka Asatani, Hiroko Yamano, Takeshi Sakaki, and Ichiro Sakata. Dense and influential core promotes daily viral information spreading in political echo chambers. *Scientific reports*, Vol. 11, p. 7491, 2021.
- [53] 山野泰子, 坂田一郎. 異質なコミュニティをつなぐノードの評価指標. WebDB Forum 2019 論文集, pp. 1–4, sep 2019.
- [54] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
- [55] Yuya Kajikawa, Yoshiyuki Takeda, Ichiro Sakata, and Katsumori Matsushima. Multiscale analysis of interfirm networks in regional clusters. *Technovation*, Vol. 30, No. 3, pp. 168–180, 2010.
- [56] Small, Trade Medium Enterprise Agency (SMEA) of the Ministry of Economy, and Industry. White paper on small and medium enterprises in Japan: Fight song for micro businesses. *White Paper on Small and Medium Enterprises in Japan*, 2014.
- [57] Marco Iansiti, Roy Levien, et al. Strategy as ecology. *Harvard business review*, Vol. 82, No. 3, pp. 68–81, 2004.
- [58] Clayton M Christensen. 繁栄のパラドクス 絶望を希望に変えるイノベーションの経済学. ハーパーコリンズ・ジャパン, 2019.
- [59] Frank Moulaert and Farid Sekia. Territorial innovation models: A critical survey. *Regional Studies*, Vol. 37, No. 3, pp. 289–302, 2003.
- [60] Ron Martin. Rebalancing the spatial economy: The challenge for regional theory. *Territory, Politics, Governance*, Vol. 3, No. 3, pp. 235–272, 2015.
- [61] Annalee Saxenian. From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China. *Studies in Comparative International Development*, Vol. 40, No. 2, pp. 35–61, 2005.

- [62] Ash Amin. Spatialities of globalisation. *Environment and Planning A*, Vol. 34, No. 3, pp. 385–399, 2001.
- [63] Eric Sheppard, Adam Tickell, Trevor J. Barnes, and Jamie Peck. Editors’ introduction: From distance to connectivity. *Reading Economic Geography*, pp. 331–335, 2004.
- [64] Ron Boschma. Proximity and innovation: A critical assessment. *Regional Studies*, Vol. 39, No. 1, pp. 61–74, 2005.
- [65] Gernot Grabher. The weakness of strong ties: the lock-in of regional development in Ruhr area. *The embedded firm. On the socioeconomics of industrial networks*, pp. 255–277, 1993.
- [66] Gernot Grabher. Ecologies of creativity: the village, the group, and the heterarchic organisation of the british advertising industry. *Environment and planning A*, Vol. 33, No. 2, pp. 351–374, 2001.
- [67] Jason Owen-Smith and Walter W. Powell. Accounting for Emergence and Novelty in Boston and Bay Area Biotechnology. *Cluster Genesis: The Emergence of Technology Clusters and Their Implications for Government Policy*, pp. 61–85, 2006.
- [68] Ronan G Powell. Modelling takeover likelihood. *Journal of Business Finance & Accounting*, Vol. 24, No. 7-8, pp. 1009–1030, 1997.
- [69] Robert Huggins and Andrew Johnston. Knowledge alliances and innovation performance: An empirical perspective on the role of network resources. *International Journal of Technology Management*, Vol. 57, No. 4, pp. 245–265, 2012.
- [70] Nicholas Bloom, Mark Schankerman, and John Van Reenen. Identifying technology spillovers and product market rivalry. *Econometrica*, Vol. 81, No. 4, pp. 1347–1393, 2013.
- [71] George S Day and Paul J H Schoemaker. Avoiding the pitfalls of emerging technologies. *California Management Review*, Vol. 42, No. 2, pp. 8–33, 2000.
- [72] George S Day and Paul J H Schoemaker. Scanning the periphery. *Harvard business review*, Vol. 83, No. 11, pp. 135–40, 142, 144–8 passim, 2005.
- [73] Maxine Robertson, Jacky Swan, and Sue Newell. The role of networks in the diffusion of technological innovation. *Journal of Management Studies*, Vol. 33, No. 3, pp. 333–359, 1996.
- [74] Elisa Giuliani and Martin Bell. The micro-determinants of meso-level learning and innovation: evidence from a chilean wine cluster. *Research policy*, Vol. 34, No. 1, pp. 47–68, 2005.

- [75] Charles A O'Reilly and Michael L Tushman. Organizational ambidexterity: Past, present, and future. *Academy of Management Perspectives*, Vol. 27, No. 4, pp. 324–338, 2013.
- [76] Manuela Presutti, Cristina Boari, Antonio Majocchi, and Xavier Molina-Morales. Distance to customers, absorptive capacity, and innovation in high-tech firms: The dark face of geographical proximity. *Journal of Small Business Management*, Vol. 57, No. 2, pp. 343–361, 2019.
- [77] Ellen Enkel and Sebastian Heil. Preparing for distant collaboration: Antecedents to potential absorptive capacity in cross-industry innovation. *Technovation*, Vol. 34, No. 4, pp. 242–260, 2014.
- [78] Katharina Kalogerakis, Christian Lüthje, and Cornelius Herstatt. Developing innovations based on analogies: experience from design and engineering consultants. *Journal of Product Innovation Management*, Vol. 27, No. 3, pp. 418–436, 2010.
- [79] Oliver Gassmann, Michael Daiber, and Ellen Enkel. The role of intermediaries in cross-industry innovation processes. *R&D Management*, Vol. 41, No. 5, pp. 457–469, 2011.
- [80] James G March. Exploration and exploitation in organizational learning. *Organization science*, Vol. 2, No. 1, pp. 71–87, 1991.
- [81] Justin JP Jansen, Frans AJ Van Den Bosch, and Henk W Volberda. Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. *Management science*, Vol. 52, No. 11, pp. 1661–1674, 2006.
- [82] Bjørn T. Asheim and Arne Isaksen. Regional innovation systems: The integration of local 'sticky' and global 'ubiquitous' knowledge. *Journal of Technology Transfer*, Vol. 27, No. 1, pp. 77–86, 2002.
- [83] Harald Bathelt, Anders Malmberg, and Peter Maskell. Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, Vol. 28, No. 1, pp. 31–56, 2004.
- [84] Alessia Sammarra and Lucio Biggiero. Heterogeneity and specificity of inter-firm knowledge flows in innovation networks. *Journal of Management Studies*, Vol. 45, No. 4, pp. 800–829, 2008.
- [85] Michael E Porter. How competitive forces shape strategy. *Harvard Business Review*, pp. 137–145, 1979.

- [86] Birger Wernerfelt. A resource-based view of the firm. *Strategic Management Journal*, Vol. 5, No. 2, pp. 171–180, 1984.
- [87] Jay B Barney. Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, Vol. 32, No. 10, pp. 1231–1241, 1986.
- [88] Jay B Barney. Firm resources and sustained competitive advantage. *Journal of management*, Vol. 17, No. 1, pp. 99–120, 1991.
- [89] Margaret A Peteraf. The cornerstones of competitive advantage: A resource-based view. *Strategic management journal*, Vol. 14, No. 3, pp. 179–191, 1993.
- [90] Coimbatore K. Prahalad and Gary Hamel. The core competence of the corporation. *Harvard business review*, Vol. 68, No. 3, pp. 79–91, 1990.
- [91] Ranjay Gulati. Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strategic management journal*, Vol. 20, No. 5, pp. 397–420, 1999.
- [92] Dovev Lavie. The competitive advantage of interconnected firms: An extension of the resource-based view. *Academy of management review*, Vol. 31, No. 3, pp. 638–658, 2006.
- [93] Toby E Stuart. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management Journal*, Vol. 21, No. 8, pp. 791–811, 2000.
- [94] Robert M Grant and Charles Baden-Fuller. A knowledge accessing theory of strategic alliances. *Journal of Management Studies*, Vol. 41, No. 1, pp. 61–84, 2004.
- [95] Marcus A Bellamy, Soumen Ghosh, and Manpreet Hora. The influence of supply network structure on firm innovation. *Journal of Operations Management*, Vol. 32, No. 6, pp. 357–373, 2014.
- [96] Walter W Powell, Kenneth W Koput, and Laurel Smith-Doerr. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative science quarterly*, pp. 116–145, 1996.
- [97] Christine M Beckman and Pamela R Haunschild. Network learning: The effects of partners' heterogeneity of experience on corporate acquisitions. *Administrative Science Quarterly*, Vol. 47, No. 1, pp. 92–124, 2002.
- [98] Brian Uzzi. Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American Sociological Review*, Vol. 64, No. 4, pp. 481–505, 1999.

- [99] Brian Uzzi. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, Vol. 61, No. 4, pp. 674–698, 1996.
- [100] Ranjay Gulati. Social structure and alliance formation patterns: A longitudinal analysis. *Administrative Science Quarterly*, Vol. 40, No. 4, pp. 619–652, 1995.
- [101] Walter Powell. Neither market nor hierarchy: Network forms of organization. *Research in Organizational Behavior*, Vol. 12, pp. 295–336, 1990.
- [102] Alan Chan, Eric Ngai, and Karen Moon. The effects of strategic and manufacturing flexibilities and supply chain agility on firm performance in the fashion industry. *European Journal of Operational Research*, Vol. 259, No. 2, pp. 486–499, 2017.
- [103] Shaw nee Vickery, Roger Calantone, and Cornelia Drge. Supply chain flexibility: An empirical study. *Journal of Supply Chain Management*, Vol. 35, No. 2, pp. 16–24, 1999.
- [104] Kevin Zhou and Fang Wu. Technological capability, strategic flexibility, and product innovation. *Strategic Management Journal*, Vol. 31, No. 5, pp. 547–561, 2010.
- [105] Thomas Y Choi, Kevin J Dooley, Manus Rungtusanatham, Thomas Y Choi, Kevin J Dooley, and Manus Rungtusanatham. Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, Vol. 19, No. 3, pp. 351–366, 2001.
- [106] Saurabh Ambulkar, Jennifer Blackhurst, and Scott Grawe. Firm’s resilience to supply chain disruptions: Scale development and empirical examination. *Journal of operations management*, Vol. 33, pp. 111–122, 2015.
- [107] Martin Christopher and Helen Peck. Building the resilient supply chain. *International Journal of Logistics Management*, Vol. 15, No. 2, pp. 1–13, 2004.
- [108] Ozgur Erol, Brian J Sauser, and Mo Mansouri. A framework for investigation into extended enterprise resilience. *Enterprise Information Systems*, Vol. 4, No. 2, pp. 111–136, 2010.
- [109] Stefano Brusoni, Andrea Prencipe, and Keith Pavitt. Knowledge specialization, organizational coupling, and the boundaries of the firm: Why do firms know more than they make? *Administrative Science Quarterly*, Vol. 46, No. 4, pp. 597–621, 2001.
- [110] Henry W Chesbrough. *Open innovation: The new imperative for creating and profiting from technology*. Harvard business school press, 2003.

- [111] Henry W Chesbrough. Open innovation: A new paradigm for understanding industrial innovation. *Oxford University Press*, No. 400, pp. 0–19, 2006.
- [112] Huong N Higgins and Judy Beckman. Abnormal returns of Japanese acquisition bidders—impact of pro-m&a legislation in the 1990s. *Pacific-Basin Finance Journal*, Vol. 14, No. 3, pp. 250–268, 2006.
- [113] Henry W Chesbrough, Wim Vanhaverbeke, and Joel West. *Open innovation: Researching a new paradigm*. Oxford University Press on Demand, 2006.
- [114] Udo Staber. The structure of networks in industrial districts. *International Journal of Urban and Regional Research*, Vol. 25, No. 3, pp. 537–552, 2001.
- [115] Michael L Tushman and Charles A O’Reilly. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review*, Vol. 38, No. 4, pp. 8–29, 1996.
- [116] Richard K Lester and Michael J Piore. *Innovation: The missing dimension*. Harvard University Press, 2009.
- [117] Robert Huggins. The success and failure of policy-implemented inter-firm network initiatives: motivations, processes and structure. *Entrepreneurship & Regional Development*, No. 2, pp. 111–135, 2000.
- [118] Gergő Tóth and Balázs Lengyel. Inter-firm inventor mobility and the role of co-inventor networks in producing high-impact innovation. *The Journal of Technology Transfer*, pp. 1–21, 2019.
- [119] Henry W Chesbrough. *Open business models: How to thrive in the new innovation landscape*. Harvard Business Press, 2006.
- [120] Jeremy Howells. Intermediation and the role of intermediaries in innovation. *Research Policy*, Vol. 35, No. 5, pp. 715–728, 2006.
- [121] Mercedes Delgado, Michael E Porter, and Scott Stern. Clusters and entrepreneurship. *Journal of economic geography*, Vol. 10, No. 4, pp. 495–518, 2010.
- [122] Michael H Best. Greater Boston’s industrial ecosystem: A manufactory of sectors. *Technovation*, Vol. 39–40, pp. 4–13, 2015.
- [123] Annalee Saxenian. *Regional Networks: Industrial Adaptation in Silicon Valley and Route 128*. Harvard University Press, Cambridge (Mass), 1994.
- [124] Jason Owen-Smith and Walter W. Powell. Knowledge networks as channels and conduits: The effects of spillovers in the boston biotechnology community. *Organization science*, Vol. 15, No. 1, pp. 5–21, 2004.

- [125] Jason Owen-Smith, Massimo Riccaboni, Fabio Pammolli, and Walter W Powell. A comparison of U.S. and european university-industry relations in the life sciences. *Management Science*, Vol. 48, No. 1, pp. 24–43, 2002.
- [126] Walter W Powell, Douglas R White, Kenneth W Koput, and Jason Owen-Smith. Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American journal of sociology*, Vol. 110, No. 4, pp. 1132–1205, 2005.
- [127] Elisa Giuliani. Network dynamics in regional clusters: Evidence from chile. *Research Policy*, Vol. 42, No. 8, pp. 1406–1419, 2013.
- [128] Anne LJ Ter Wal and Ron Boschma. Co-evolution of firms, industries and networks in space. *Regional Studies*, Vol. 45, No. 7, pp. 919–933, 2011.
- [129] Olivier Crevoisier and Hugues Jeannerat. Territorial knowledge dynamics: from the proximity paradigm to multi-location milieus. *European planning studies*, Vol. 17, No. 8, pp. 1223–1241, 2009.
- [130] Laura James, Geert Vissers, Anders Larsson, and Margareta Dahlström. Territorial knowledge dynamics and knowledge anchoring through localized networks: The automotive sector in västra götaland. *Regional studies*, Vol. 50, No. 2, pp. 233–244, 2016.
- [131] Marco Iansiti and Kim B Clark. Integration and dynamic capability: Evidence from product development in automobiles and mainframe computers. *Industrial and corporate change*, Vol. 3, No. 3, pp. 557–605, 1994.
- [132] Ron Martin and Peter Sunley. Conceptualizing cluster evolution: Beyond the life cycle model? *Regional Studies*, Vol. 45, No. 10, pp. 1299–1318, 2011.
- [133] Frederic Edward Clements. *Plant succession: An analysis of the development of vegetation*. No. 242. Carnegie Institution of Washington, 1916.
- [134] 嶋田正和, 山村則男, 粕谷英一, 伊藤嘉昭. 動物生態学新版. 海游舎, 東京, 2005.
- [135] Kozo Kiyota and Tetsuji Okazaki. Industrial policy cuts two ways: Evidence from cotton-spinning firms in japan, 1956–1964. *The Journal of Law and Economics*, Vol. 53, No. 3, pp. 587–609, 2010.
- [136] ヨーゼフ・シュンペーター, 塩野谷祐一, 東畑精一, 中山伊知郎. 経済発展の理論. 塩野谷裕一ほか訳, 岩波文庫, 1977.
- [137] 村上泰亮. 反古典の政治経済学要綱—来世紀のための覚書. 中央公論社, 1994.

- [138] Eva Bianconi, Allison Piovesan, Federica Facchin, Alina Beraudi, Raffaella Casadei, Flavia Frabetti, Lorenza Vitale, Maria Chiara Pelleri, Simone Tassani, Francesco Piva, et al. An estimation of the number of cells in the human body. *Annals of human biology*, Vol. 40, No. 6, pp. 463–471, 2013.
- [139] 坂田一郎, 梶川裕矢, 武田善行, 橋本正洋, 柴田尚樹, 松島克守. 地域クラスターのネットワーク形成のダイナミクス. *Research Institute of Economy, Trade and Industry (RIETI) Discussion Paper Series*, Vol. 7, pp. 1–41, 2007.
- [140] Annalee Saxenian. The origins and dynamics of production networks in silicon valley. *Research Policy*, Vol. 20, No. 5, pp. 423–437, 1991.
- [141] Kai Foerstl, Carsten Reuter, Evi Hartmann, and Constantin Blome. Managing supplier sustainability risks in a dynamically changing environment—sustainable supplier management in the chemical industry. *Journal of Purchasing and Supply Management*, Vol. 16, No. 2, pp. 118–130, 2010.
- [142] Ola Laaksonen and Mirva Peltoniemi. The essence of dynamic capabilities and their measurement. *International Journal of Management Reviews*, Vol. 20, No. 2, pp. 184–205, 2018.
- [143] Danilo Brozovic. Strategic flexibility A review of the literature. *International Journal of Management Reviews*, Vol. 20, No. 1, pp. 3–31, 2018.
- [144] Giulio Rossetti and Rmy Cazabet. Community discovery in dynamic networks: A survey. *ACM Computing Surveys*, Vol. 51, No. 2, pp. 1–37, 2018.
- [145] Yoshiyuki Takeda, Yuya Kajikawa, Ichiro Sakata, and Katsumori Matsushima. An analysis of geographical agglomeration and modularized industrial networks in a regional cluster: A case study at yamagata prefecture in japan. *Technovation*, Vol. 28, No. 8, pp. 531–539, 2008.
- [146] Yuya Kajikawa, Junichiro Mori, and Ichiro Sakata. Identifying and bridging networks in regional clusters. *Technological Forecasting and Social Change*, Vol. 79, No. 2, pp. 252–262, 2012.
- [147] Junichiro Mori, Yuya Kajikawa, Hisashi Kashima, and Ichiro Sakata. Machine learning approach for finding business partners and building reciprocal relationships. *Expert Systems with Applications*, Vol. 39, No. 12, pp. 10402–10407, 2012.
- [148] Hirokazu Kawamoto, Hideki Takayasu, and Misako Takayasu. Network anatomy controlling abrupt-like percolation transition. *Scientific Reports*, Vol. 7, No. 1, p. 163, 2017.

- [149] Hayato Goto, Eduardo Viegas, Henrik Jensen, Hideki Takayasu, and Misako Takayasu. Appearance of unstable monopoly state caused by selective and concentrative mergers in business networks. *Scientific Reports*, Vol. 7, No. 1, p. 5064, 2017.
- [150] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pp. 35–41, 1977.
- [151] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, Vol. 1, No. 3, pp. 215–239, 1978.
- [152] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, Vol. 69, No. 2, p. 026113, 2004.
- [153] Mark EJ Newman. Fast algorithm for detecting community structure in networks. *Physical review E*, Vol. 69, No. 6, p. 066133, 2004.
- [154] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, Vol. 18, No. 1, pp. 39–43, 1953.
- [155] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.
- [156] Francois Lorrain and Harrison C White. Structural equivalence of individuals in social networks. *The Journal of mathematical sociology*, Vol. 1, No. 1, pp. 49–80, 1971.
- [157] Roger Guimera, Marta Sales-Pardo, and Luis AN Amaral. Classes of complex networks defined by role-to-role connectivity profiles. *Nature physics*, Vol. 3, No. 1, pp. 63–69, 2007.
- [158] Gergely Palla, Albert-László Barabási, and Tamás Vicsek. Quantifying social group evolution. *Nature*, Vol. 446, No. 7136, pp. 664–667, 2007.
- [159] Piotr Bródka, Stanisaw Saganowski, and Przemysaw Kazienko. Ged: the method for group evolution discovery in social networks. *Social Network Analysis and Mining*, Vol. 3, No. 1, pp. 1–14, 2013.
- [160] Sitaram Asur, Srinivasan Parthasarathy, and Duygu Ucar. An event-based framework for characterizing the evolutionary behavior of interaction graphs. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, p. 913921, New York, NY, USA, 2007. Association for Computing Machinery.

- [161] Mansoureh Takaffoli, Farzad Sangi, Justin Fagnan, and Osmar R. Zaiane. Modec — modeling and detecting evolutions of communities. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011.
- [162] Piotr Bródka, Katarzyna Musiał, and Przemysław Kazienko. A performance of centrality calculation in social networks. *2009 International Conference on Computational Aspects of Social Networks*, pp. 24–31, 2009.
- [163] Stanisław Saganowski, Piotr Bródka, and Przemysław Kazienko. Influence of the user importance measure on the group evolution discovery. *Foundations of Computing and Decision Sciences*, Vol. 37, No. 4, pp. 293–303, 2012.
- [164] Ziwei He, Etienne Gael Tajeuna, Shengrui Wang, and Mohamed Bouguessa. A comparative study of different approaches for tracking communities in evolving social networks. *arXiv*, 2019.
- [165] Mansoureh Takaffoli, Reihaneh Rabbany, and Osmar R Zaiane. Community evolution prediction in dynamic social networks. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pp. 9–16, 2014.
- [166] Hiroko Yamano, Hajime Sasaki, and Ichiro Sakata. Metabolism of inter-firm transactions in regional networks. *Proceedings of the 2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, pp. 1–8, 2017.
- [167] Hiroko Yamano and Ichiro Sakata. Regional differences and commonalities of inter-firm transaction metabolism. *Proceedings of the 2017 International Society for Professional Innovation Management (ISPIM) Innovation Symposium*, pp. 1–13, 2017.
- [168] Hiroko Yamano and Ichiro Sakata. Latent pattern extraction and factorization of firm bankruptcies and metabolism in japan. *Proceedings of the 2019 Portland International Conference on Management of Engineering and Technology (PICMET)*, pp. 1–12, 2019.
- [169] Dieter Rasch, Klaus D Kubinger, and Karl Moder. The two-sample t test: pre-testing its assumptions does not pay off. *Statistical papers*, Vol. 52, No. 1, pp. 219–231, 2011.
- [170] Yukiko Saito. The impact of the great east japan earthquake on companies in the non-affected areas: Structure of the inter-company network of supply chains and its implication (Japanese). Discussion papers (japanese), Research Institute of Economy, Trade and Industry (RIETI), 2012.
- [171] Duncan J Watts. *Small worlds: the dynamics of networks between order and randomness*, Vol. 9. Princeton university press, 2004.

- [172] Hiroko Yamano, Kimitaka Asatani, and Ichiro Sakata. Evaluating nodes of latent mediators in heterogeneous communities. *Scientific reports*, Vol. 10, No. 1, p. 8456, 2020.
- [173] Hiroko Yamano, Kimitaka Asatani, and Ichiro Sakata. Evaluating nodes of latent mediators in heterogeneous communities. *Book of Abstracts of the 8th International Conference on Complex Networks and their Applications (Complex Networks 2019)*, pp. 90–92, 2019.
- [174] Bowen Yan and Steve Gregory. Finding missing edges in networks based on their community structure. *Physical Review E*, Vol. 85, No. 5, p. 056112, 2012.
- [175] Carlo Vittorio Cannistraci, Gregorio Alanis-Lobato, and Timothy Ravasi. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Scientific reports*, Vol. 3, p. 1613, 2013.
- [176] Linyuan Lü, Liming Pan, Tao Zhou, Yi-Cheng Zhang, and H Eugene Stanley. Toward link predictability of complex networks. *Proceedings of the National Academy of Sciences*, Vol. 112, No. 8, pp. 2325–2330, 2015.
- [177] Jingyi Ding, Licheng Jiao, Jianshe Wu, and Fang Liu. Prediction of missing links based on community relevance and ruler inference. *Knowledge-Based Systems*, Vol. 98, pp. 200–215, 2016.
- [178] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, Vol. 286, No. 5439, pp. 509–512, 1999.
- [179] Lada A Adamic and Eytan Adar. Friends and neighbors on the web. *Social networks*, Vol. 25, No. 3, pp. 211–230, 2003.
- [180] Gourab Ghoshal and Albert-László Barabási. Ranking stability and super-stable nodes in complex networks. *Nature communications*, Vol. 2, p. 394, 2011.
- [181] Jure Leskovec and Julian J McAuley. Learning to discover social circles in ego networks. In *Advances in neural information processing systems*, pp. 539–547, 2012.
- [182] Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, Vol. 21, No. 11, pp. 1129–1164, 1991.
- [183] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *Nature*, Vol. 406, No. 6794, pp. 378–382, 2000.
- [184] Hiroko Yamano and Ichiro Sakata. Role of brokers and clusters in the inter-firm network dynamics: Evolution map perspective. *Proceedings of the 2022 Portland International Conference on Management of Engineering and Technology (PICMET)*, forthcoming, pp. 1–10, 2022.

- [185] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, Vol. 2008, No. 10, p. P10008, 2008.
- [186] R Lambiotte, J C Delvenne, and M Barahona. Laplacian dynamics and multiscale modular structure in networks. *arXiv*, 2009.
- [187] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, Vol. 27, No. 3, pp. 379–423, 1948.
- [188] In-Kwon Yeo and Richard A Johnson. A new family of power transformations to improve normality or symmetry. *Biometrika*, Vol. 87, No. 4, pp. 954–959, 2000.
- [189] Max Kuhn and Davis Vaughan. *parsnip: A Common API to Modeling and Analysis Functions*, 2020. R package version 0.1.3.
- [190] Brandon M. Greenwell and Bradley C. Boehmke. Variable importance plots—an introduction to the vip package. *The R Journal*, Vol. 12, No. 1, pp. 343–366, 2020.
- [191] Georgios Diakidis, Despoina Karna, Dimitris Fasarakis-Hilliard, Dimitrios Vogiatzis, and George Paliouras. Predicting the evolution of communities in social networks. In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics (WIMS '15)*, pp. 1–6, 2015.
- [192] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics Reports*, Vol. 716–717, pp. 1–58, 2017.
- [193] Ivo Mossig and Lars Schieber. Driving forces of cluster evolution – growth and lock-in of two german packaging machinery clusters. *European Urban and Regional Studies*, Vol. 23, No. 4, pp. 594–611, 2016.
- [194] Denis Maillat, Bruno Lecoq, Florian Nemeti, and Marc Pfister. Technology district and innovation: The case of the swiss jura arc. *Regional Studies*, Vol. 29, No. 3, pp. 251–263, 1995.
- [195] Harald Bathelt and Nina Schuldt. Between luminaires and meat grinders: International trade fairs as temporary clusters. *Regional Studies*, Vol. 42, No. 6, pp. 853–868, 2008.
- [196] Elisa Giuliani. Role of technological gatekeepers in the growth of industrial clusters: Evidence from chile. *Regional Studies*, Vol. 45, No. 10, pp. 1329–1348, 2011.
- [197] Elisa Giuliani. Networks and heterogeneous performance of cluster firms. *Applied evolutionary economics and economic geography*, pp. 161–179, 2007.

- [198] Hiroshi Okamoto and Xu-le Qiu. Modular decomposition of markov chain: detecting hierarchical organization of pervasive communities. *arXiv preprint arXiv:1909.07066*, 2019.
- [199] 岡本洋. 全脳ネットワーク分析: 複雑ネットワーク科学とコネクトームとの融合から全脳アーキテクチャにせまる. *知能と情報*, Vol. 28, No. 2, pp. 28–37, 2016.
- [200] Daryl J Daley and David Kendall. Epidemics and rumours. *Nature*, Vol. 204, No. 4963, p. 1118, 1964.
- [201] Jessica T Davis, Nicola Perra, Qian Zhang, Yamir Moreno, and Alessandro Vespignani. Phase transitions in information spreading on structured populations. *Nature Physics*, Vol. 16, No. 5, pp. 590–596, 2020.
- [202] Zakariya Ghalmane, Chantal Cherifi, Hocine Cherifi, and Mohammed Hassouni. Centrality in complex networks with overlapping community structure. *Scientific Reports*, Vol. 9, No. 1, p. 10133, 2019.
- [203] Eri Yamada and Tetsu Kawakami. Distribution of industrial growth in nagoya metropolitan area, japan: An exploratory analysis using geographical and technological proximities. *Regional Studies*, Vol. 50, No. 11, pp. 1–13, 2015.

付録 A

A.1 第3章の補足資料

地方別新陳代謝度の分布

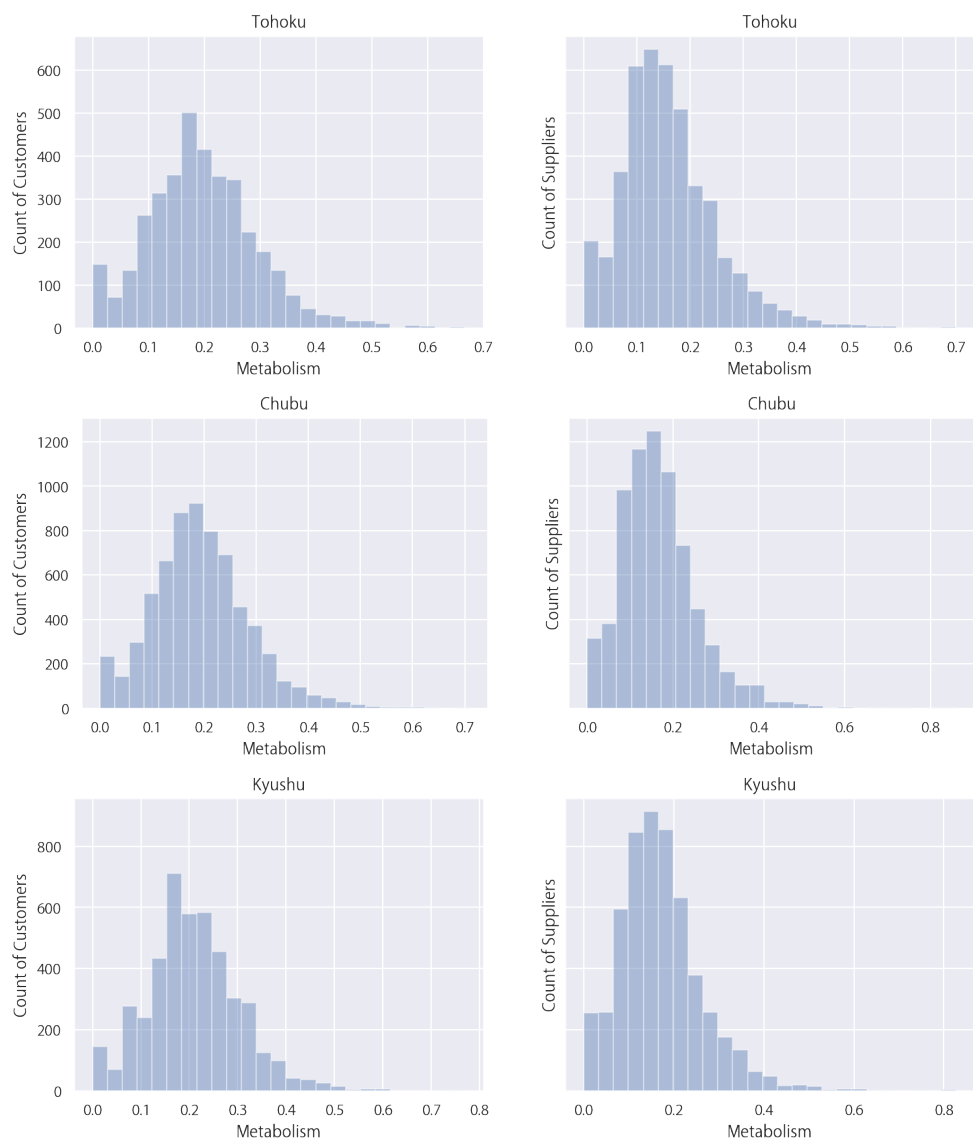


図 A.1: 地方別新陳代謝度の分布（左は発注社企業，右は受注社企業）

新陳代謝度の多重比較検定 1

(1) 取引先を替えない企業群の有意差検定（全企業対象）

新陳代謝度が0で取引先を全く替えなかった企業群と、取引先を少なくとも1社は変更する企業群との独立2群間において存続年数の代表値の有意差検定を行った。対象企業は、10年間の期間中に存在した東北地方の全企業155,883社（東北地方の企業と取引関係のある他地域の企業を含む）のうち、2年以上存続し新陳代謝度が計算できた企業140,251社。

表 A.1: 独立2群間の有意差検定（全企業）

条件式	企業数	平均存続年数
$metabolism > 0$	73,015	8.551
$metabolism = 0$	67,236	5.807
Welch 検定	$t = 194.04$	$p\text{-value} < 2.2e-16$
Mann-Whitney 検定	$statistic = 1172729987$	$p\text{-value} < 2.2e-16$

(2) 新陳代謝度の分位点別企業群における存続年数の有意差検定（全企業対象）

新陳代謝度の値の分位点で企業群を分類し、各群の存在期間に有意差があるかどうかを検定した。なお、全企業を対象にしたデータでは、新陳代謝度が0の企業が半数近くに及ぶため、25分位点より小さい分位点を作ることはできない。その一方、新陳代謝度が高い企業は分散が大きいため、分位点は四分位点に加え、95.0、97.5の分位点を追加した。各分位点における企業数と平均新陳代謝度を表 A.2 に示す。

表 A.2: 分位点別企業数と平均新陳代謝度（全企業）

分位点	企業数	平均新陳代謝度
25	69,275	0.001
50	35,954	0.077
75	21,063	0.174
90	6,897	0.275
95	3,639	0.366
97.5	3,304	0.550

図 A.2 は、分位点水準での各企業群における存続年数の分布と Tukey-Kramer 法による群間有意差の検定結果である。図の上端のアルファベットの昇順は平均の昇順に対応し、平均間に有意差がない場合には同じアルファベットが付与される。今回は全企業群において存続年数の平均値に有意差があり、アルファベットの重複は見られなかった。また、新陳代謝度の分位点が97.5の範囲にいる企業群の存続年数は分散が大きく、平均存続年数が最も短かった。分位点25範囲の企業群はその次に平均存続年数が短く、分散も大きかった。分位

点 50 から分位点 97.5 までは新陳代謝度が下がるにつれて存続年数が長くなる傾向があり、全企業群中、最も存続年数が長いのは分位点 50 の企業群で、その平均新陳代謝度は 0.077 であった。この結果は新陳代謝度は高すぎても低すぎても存続年数が短くなるというこれまでの分析結果と一致している。

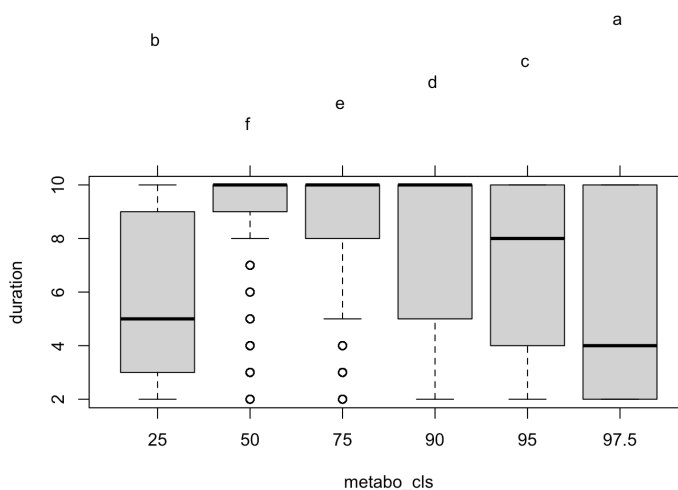


図 A.2: 新陳代謝度の分位点水準での存続年数の分布と TK 検定（全企業）

群間有意差の 95%信頼区間による平均の差を図 A.3 に示す。今回は 6 群の多重比較検定なので、 $6 \times 5/2 = 15$ 通りの組み合わせの 2 群について、差の推定値は \cdot で、95%信頼区間の下限と上限は (-) で示されている。点線は平均の差がゼロであることを意味するため、95%信頼区間が点線を跨がないことは危険率 5% の水準で有意であることを意味する。

Tukey-Kramer 検定の結果、いずれの群間においても点線を跨ぐ信頼区間は見られなかった。また、最も代表値の差が小さい群は分位点 97.5 と分位点 25 にいる企業群であり、新陳代謝度が高すぎても低すぎても存続年数が短くなるという仮説が支持される結果となった。平均新陳代謝度が 0.077 および 0.174 である分位点 75 と分位点 50 の企業群の差も小さく、1~2 割程度の取引を入れ替えることが企業が長く存続するために必要であることが伺えた。最も代表値の差が大きい企業群は分位点 50 と分位点 25、および分位点 97.5 と分位点 50 であった。新陳代謝度は両極端ではなく、中庸であることが長期間存続するために必要であることが示唆される。

6 つの分位点で分けられた企業群のすべての組み合わせにおける統計量を表 A.3 に示す。いずれの組み合わせにおいても、 $P\text{-value} = 2e-16$ となり、帰無仮説は棄却され、群間に有意差があることが認められた。

95% family-wise confidence level

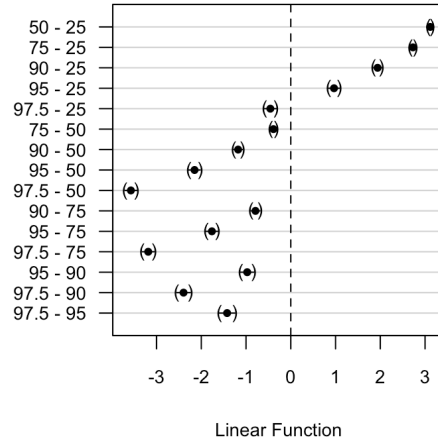


図 A.3: 新陳代謝度の分位点水準での存続年数の 95%信頼区間 (全企業)

表 A.3: 新陳代謝度の分位点水準での存続年数の有意差検定

Null hypothesis	Estimate	Std. Error	t value	Pr(> t)
50 - 25 == 0	3.116	0.017	186.161	<2e-16 ***
75 - 25 == 0	2.729	0.020	134.664	<2e-16 ***
90 - 25 == 0	1.939	0.033	59.646	<2e-16 ***
95 - 25 == 0	0.966	0.044	22.059	<2e-16 ***
97.5 - 25 == 0	-0.457	0.046	-9.955	<2e-16 ***
75 - 50 == 0	-0.387	0.022	-17.338	<2e-16 ***
90 - 50 == 0	-1.177	0.034	-34.758	<2e-16 ***
95 - 50 == 0	-2.150	0.045	-47.993	<2e-16 ***
97.5 - 50 == 0	-3.573	0.047	-76.314	<2e-16 ***
90 - 75 == 0	-0.789	0.036	-22.091	<2e-16 ***
95 - 75 == 0	-1.763	0.046	-38.126	<2e-16 ***
97.5 - 75 == 0	-3.185	0.048	-66.100	<2e-16 ***
95 - 90 == 0	-0.973	0.053	-18.447	<2e-16 ***
97.5 - 90 == 0	-2.396	0.054	-43.974	<2e-16 ***
97.5 - 95 == 0	-1.423	0.062	-22.989	<2e-16 ***
Signif. codes	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'

(3) 存続年数別の企業群における新陳代謝度の有意差検定（全企業対象）

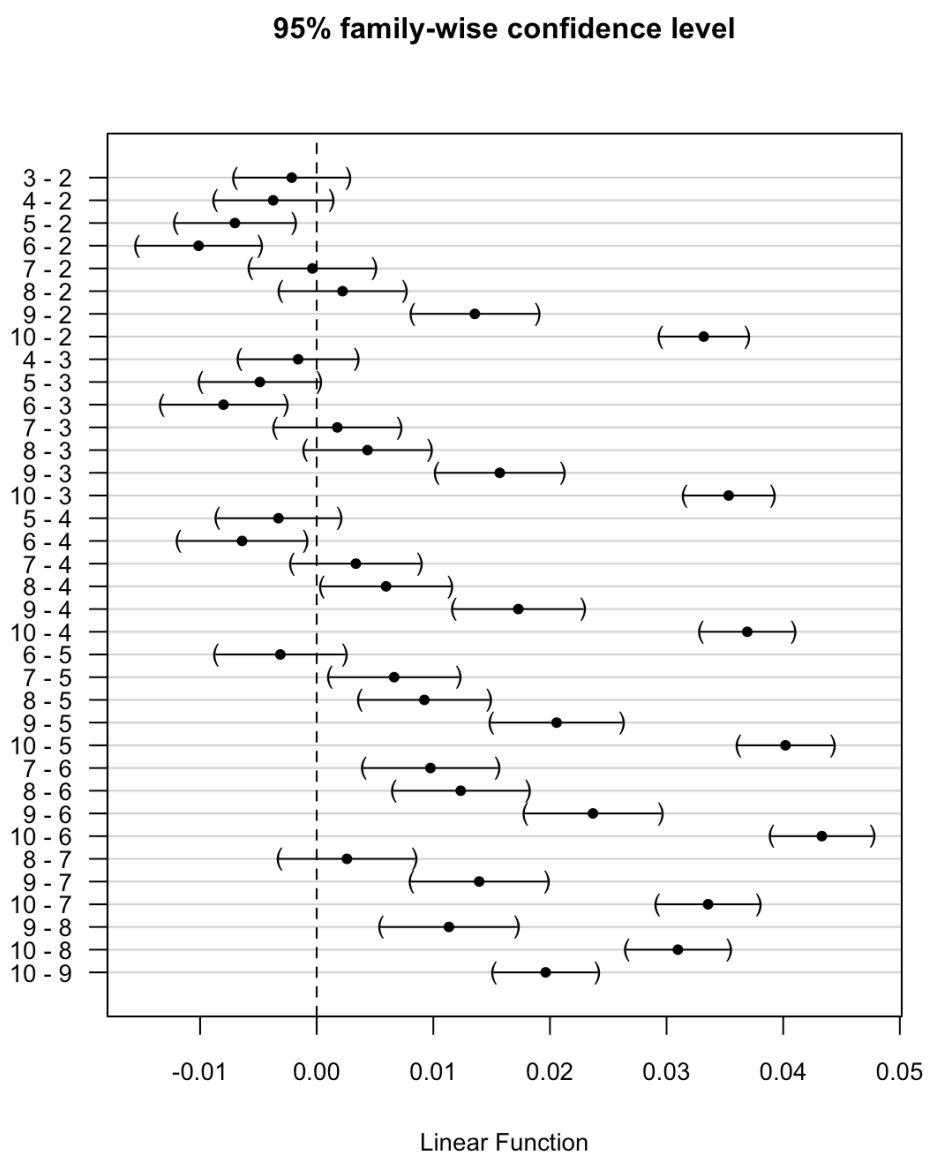


図 A.4: 存続年数別企業群における新陳代謝度の 95%信頼区間

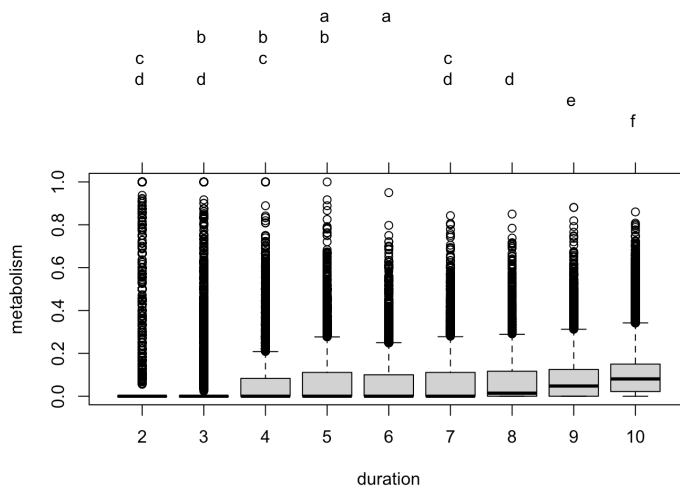


図 A.5: 存続年数別企業群における新陳代謝度の有意差検定

(4) 新陳代謝度の分位点別企業群における企業年齢の有意差検定（全企業対象）

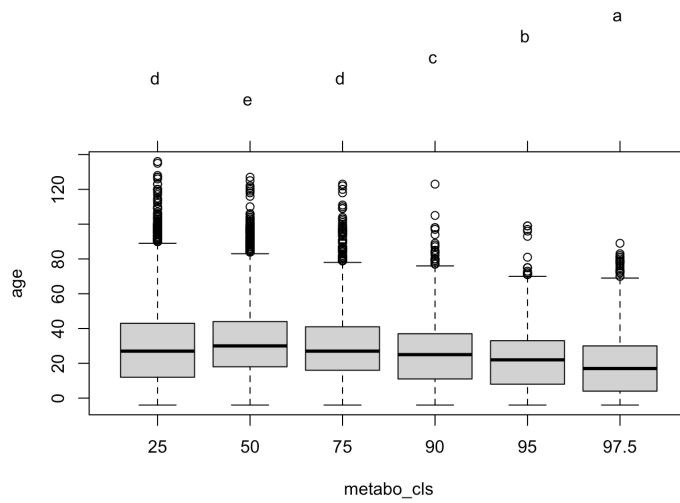


図 A.6: 新陳代謝度分位点における企業年齢の多重比較検定

95% family-wise confidence level

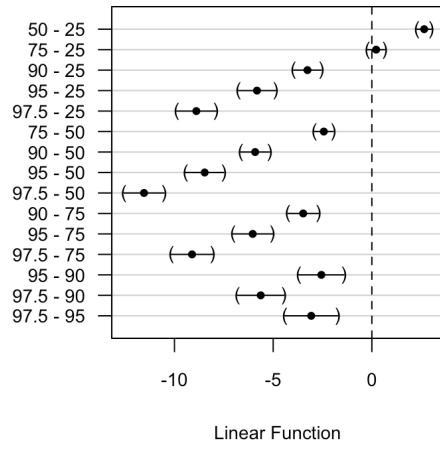


図 A.7: 新陳代謝度分位点における企業年齢の群間比較の95%信頼区間

(5) 企業年齢の分位点別企業群における新陳代謝度の有意差検定 (全企業対象)

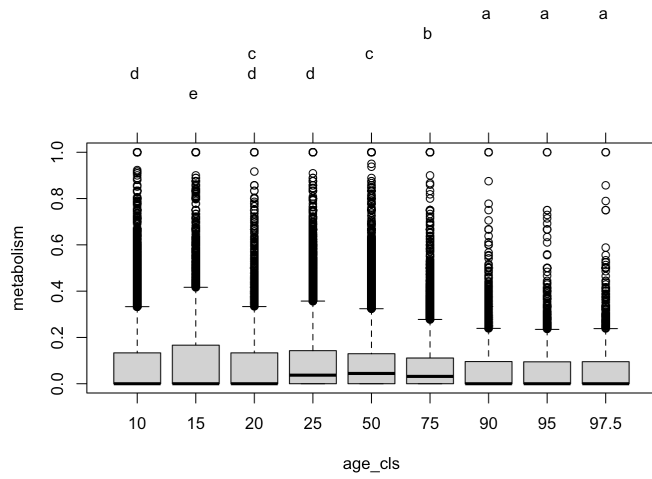


図 A.8: 企業年齢分位点における新陳代謝度の多重比較検定

95% family-wise confidence level

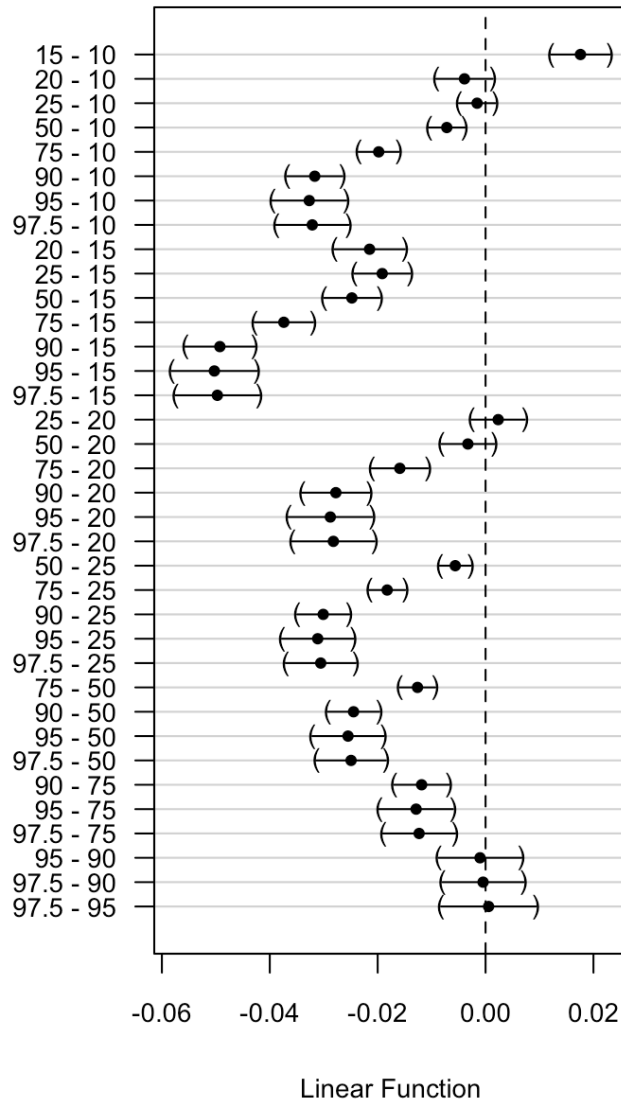


図 A.9: 企業年齢分位点における新陳代謝度の群間比較の 95%信頼区間

不等分散でノンパラメトリックな多重比較検定

Tukey-Kramer 法は多重比較検定で最も多く使われる検定手法ではあるが、正規性と等分散性が仮定されている。これに代わる方法として、分散が等しくないときは Games-Howell 法、ノンパラメトリックである場合は Steel-Dwass 法が推奨されている。以下では、新陳代謝度の分位点水準における企業群の存続年数について、これらの手法を用いた検定を行った。対象データは第 3 章の多重比較検定で用いたデータと同じ、新陳代謝度が計算できた取引数 10 以上の東北地方の企業である。

検定の結果、Games-Howell 法、Steel-Dwass 法は Tukey-Kramer 法に比べると群間有意差が小さくなる傾向があるものの、全体的な傾向として、分位点 2.5 および分位点 97.5 の企業群が他の全ての分位点の企業群と存続年数に有意差があることは変わらなかった。新陳代謝度が高すぎても低すぎても存続年数が短くなる傾向がここでも確認されたと言える。

表 A.4: Pairwise comparisons using Games-Howell test

	2.5	5	10	25	50	75	90	95
5	< 2e-16	-	-	-	-	-	-	-
10	2.0e-10	1.00000	-	-	-	-	-	-
25	4.9e-10	0.67373	0.17990	-	-	-	-	-
50	4.7e-10	0.15146	0.00376	0.88452	-	-	-	-
75	< 2e-16	1.00000	1.00000	0.25964	0.00894	-	-	-
90	< 2e-16	0.33483	0.23554	0.00220	0.00013	0.23400	-	-
95	4.7e-10	0.00031	0.00017	2.5e-06	3.4e-07	0.00017	0.11165	-
97.5	1.9e-11	2.2e-12	8.0e-13	2.0e-13	2.9e-13	1.0e-12	1.3e-08	0.01799

P value adjustment method: none — alternative hypothesis: two.sided

表 A.5: Pairwise comparisons using Dwass-Steele-Critchlow-Fligner all-pairs test

	2.5	5	10	25	50	75	90	95
5	< 2e-16	-	-	-	-	-	-	-
10	< 2e-16	0.84786	-	-	-	-	-	-
25	< 2e-16	0.00053	0.00811	-	-	-	-	-
50	< 2e-16	1.9e-06	1.2e-05	0.78168	-	-	-	-
75	< 2e-16	0.16368	0.84948	0.65016	0.02963	-	-	-
90	< 2e-16	1.00000	0.93758	0.00501	5.2e-05	0.33642	-	-
95	< 2e-16	0.14990	0.00098	1.2e-08	2.8e-11	2.4e-05	0.17323	-
97.5	8.7e-14	1.4e-09	6.5e-14	< 2e-16	< 2e-16	6.5e-14	5.5e-09	0.03396

P value adjustment method: single-step

新陳代謝度の多重比較検定 2

帝国データバンク (TDB) の企業間取引データは、取引関係のある「発注社企業」と「受注社企業」がセットになっており、取引関係の方向は1種類である。これに対し、東京商工リサーチ (TSR) の場合、調査を受けた会社毎に取引関係データが格納されており、ある企業があげた取引先企業は「仕入先」「販売先」「株主」のいずれかに分類されている。このため、TSR データの取引関係には TDB データの「発注社・受注社」に相当する「仕入先・販売先」に加え、企業が答えたデータなのか、取引先企業に挙げられたデータなのかという2種類の方向が存在する。

本論文で定義した新陳代謝度は、調査を受けた企業が、主要取引先としてあげる企業群を年次推移に伴い変化させているかを計算している。これは取引関係を片側から捉えるものであり、ある企業が把握している主要な取引関係の変動に相当する。しかし、企業間取引をネットワークとして考えた場合、仮にある企業が主要取引先を10年間変更しなくても、他社から主要取引先として挙げられる回数が増えれば、ネットワーク上の取引関係は変動する。たとえば、東京の大企業にとっては主要取引先でなくても、東北地方の企業には主要取引先ということがありうる。そこで本節では、TSR データを用いて、主要取引先として挙げられた方のカウントも含めた場合の新陳代謝度（ここでは新陳代謝度2とする）を計算した。この場合の平均取引数は各企業の存続期間における平均次数であり、存続年数は各企業が対象ネットワーク内にエッジを少なくとも1本は保持している期間である。

平均取引数が10社以上の企業19,151社について、新たに計算した新陳代謝度2と存続年数の分布をプロットしたのが図A.10である。新陳代謝度2の方が分布傾向が顕著（存続年数が短い企業の新陳代謝度の分散が大きい）であり、片側から見た場合よりも、新陳代謝度が高すぎる企業の存続がより難しいことが示されている。

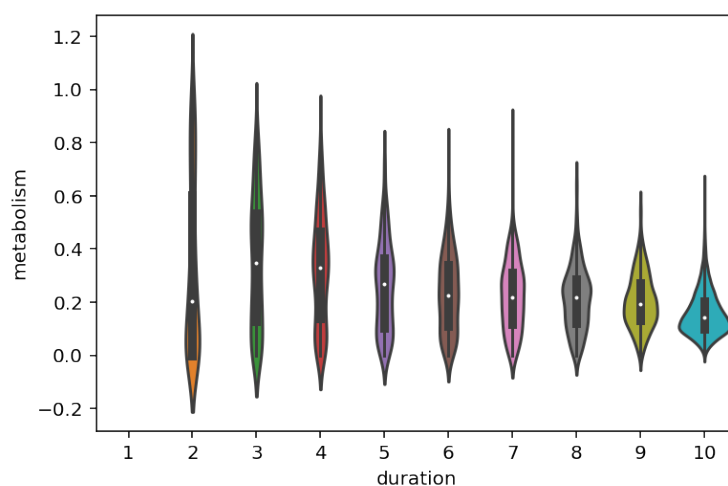


図 A.10: 東北地方の新陳代謝度2と存続年数の分布

第3章と同様に、新陳代謝度2の分位点別企業群の多重比較検定を行なった。今回は（分位点0の企業群と分位点2.5の企業群の平均新陳代謝度が異なっていたため、分位点0を含む、10群での多重比較検定となった。分位点別企業数と基本統計量を表A.6に示す。

表 A.6: 新陳代謝度2の分位点別企業数と基本統計量

分位点	企業数	平均値	中央値	標準偏差
0	473	0.0125	0.0130	0.0112
2.5	474	0.0392	0.0397	0.00508
5	947	0.0574	0.0575	0.00578
10	2840	0.0844	0.0851	0.00926
25	4734	0.123	0.122	0.0139
50	4734	0.180	0.179	0.0208
75	2841	0.257	0.254	0.0236
90	947	0.331	0.330	0.0176
95	473	0.401	0.4	0.0215
97.5	473	0.592	0.552	0.133
NA	215	NA	NA	NA

図A.11はTukey-Kramer法による多重比較検定結果である。存続年数が最も短いのは新陳代謝度が最も小さい分位点97.5の企業群であり、新陳代謝度が最も大きい分位点0、分位点95の企業群がこれに続いている。分位点0の企業群は平均新陳代謝度が0.0125で、大半が取引先を全く変えなかった企業であり、分位点97.5の企業群は平均新陳代謝度が0.592であることから、6割がた取引先を入れ替えた企業である。存続年数が最も長いのは分位点25の企業群で平均新陳代謝度は0.123であり、第3章とほぼ一致する分析結果であった。

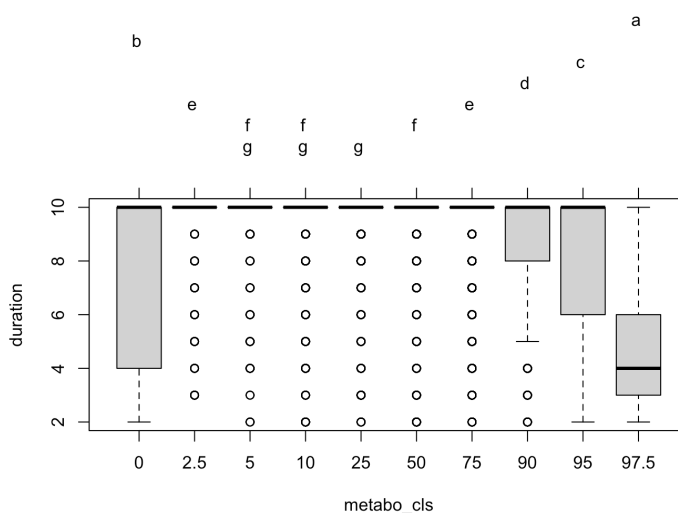


図 A.11: 新陳代謝度2の分位点と存続年数の多重比較検定結果

10個の分位点で分けられた企業群のすべての組み合わせにおける統計量を表A.7に、平均存続年数の群間有意差の95%信頼区間によるプロットを図A.12に示す。

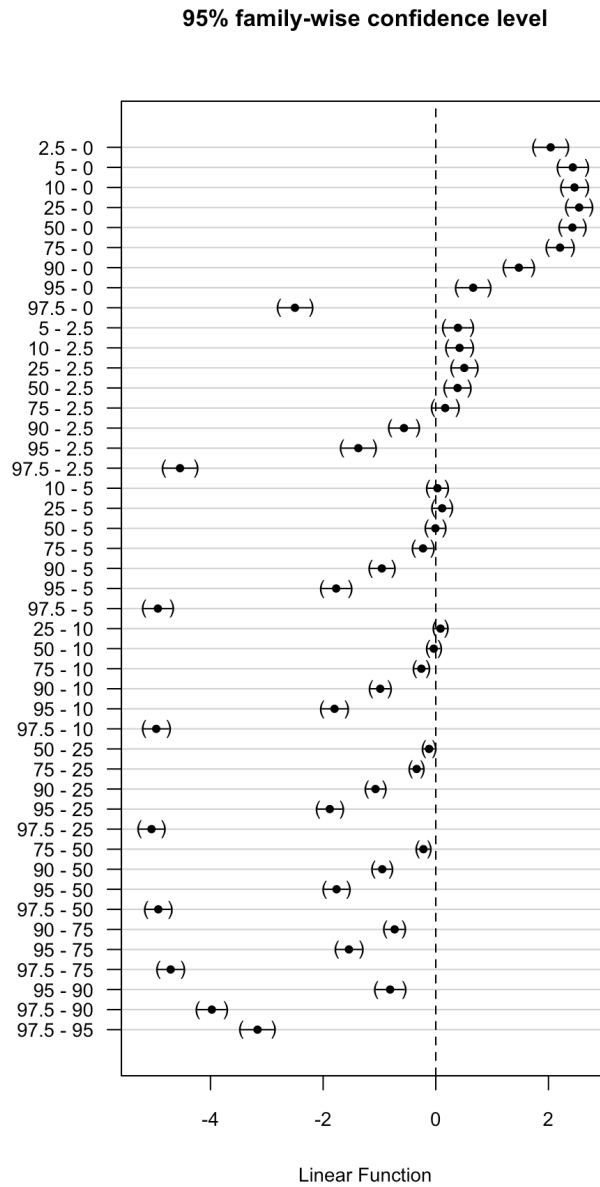


図 A.12: 新陳代謝度 2 の分位水準での存続年数の 95%信頼区間

表 A.7: 新陳代謝度 2 の分位点水準での多重比較検定

Null hypothesis	Estimate Std.	Error	t value	Pr(> t)
2.5 - 0 == 0	2.040	0.093	21.981	<0.01 ***
5 - 0 == 0	2.433	0.080	30.263	<0.01 ***
10 - 0 == 0	2.462	0.071	34.722	<0.01 ***
25 - 0 == 0	2.546	0.069	36.976	<0.01 ***
50 - 0 == 0	2.426	0.069	35.233	<0.01 ***
75 - 0 == 0	2.206	0.071	31.115	<0.01 ***
90 - 0 == 0	1.475	0.080	18.348	<0.01 ***
95 - 0 == 0	0.664	0.093	7.151	<0.01 ***
97.5 - 0 == 0	-2.501	0.093	-26.941	<0.01 ***
5 - 2.5 == 0	0.393	0.080	4.894	<0.01 ***
10 - 2.5 == 0	0.422	0.071	5.961	<0.01 ***
25 - 2.5 == 0	0.506	0.069	7.358	<0.01 ***
50 - 2.5 == 0	0.386	0.069	5.614	<0.01 ***
75 - 2.5 == 0	0.166	0.071	2.350	0.320
90 - 2.5 == 0	-0.565	0.080	-7.029	<0.01 ***
95 - 2.5 == 0	-1.376	0.093	-14.826	<0.01 ***
97.5 - 2.5 == 0	-4.541	0.093	-48.936	<0.01 ***
10 - 5 == 0	0.029	0.054	0.544	1.000
25 - 5 == 0	0.113	0.051	2.222	0.402
50 - 5 == 0	-0.007	0.051	-0.138	1.000
75 - 5 == 0	-0.227	0.054	-4.232	<0.01 ***
90 - 5 == 0	-0.958	0.066	-14.598	<0.01 ***
95 - 5 == 0	-1.769	0.080	-22.005	<0.01 ***
97.5 - 5 == 0	-4.934	0.080	-61.378	<0.01 ***
25 - 10 == 0	0.084	0.034	2.474	0.250
50 - 10 == 0	-0.036	0.034	-1.067	0.985
75 - 10 == 0	-0.256	0.038	-6.752	<0.01 ***
90 - 10 == 0	-0.987	0.054	-18.422	<0.01 ***
95 - 10 == 0	-1.798	0.071	-25.359	<0.01 ***
97.5 - 10 == 0	-4.963	0.071	-69.998	<0.01 ***
50 - 25 == 0	-0.120	0.029	-4.089	<0.01 **
75 - 25 == 0	-0.340	0.034	-10.024	<0.01 ***
90 - 25 == 0	-1.071	0.051	-21.068	<0.01 ***
95 - 25 == 0	-1.882	0.069	-27.333	<0.01 ***
97.5 - 25 == 0	-5.047	0.069	-73.304	<0.01 ***
75 - 50 == 0	-0.220	0.034	-6.483	<0.01 ***
90 - 50 == 0	-0.951	0.051	-18.707	<0.01 ***
95 - 50 == 0	-1.762	0.069	-25.590	<0.01 ***
97.5 - 50 == 0	-4.927	0.069	-71.561	<0.01 ***
90 - 75 == 0	-0.731	0.054	-13.647	<0.01 ***
95 - 75 == 0	-1.542	0.071	-21.751	<0.01 ***
97.5 - 75 == 0	-4.707	0.071	-66.391	<0.01 ***
95 - 90 == 0	-0.811	0.080	-10.090	<0.01 ***
97.5 - 90 == 0	-3.976	0.080	-49.463	<0.01 ***
97.5 - 95 == 0	-3.165	0.093	-34.092	<0.01 ***
Signif. codes	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'

A.2 第4章の補足資料



図 A.13: Facebook 10 ego users' network with 15 communities. The left and right graphs represent the same combined Facebook network with different node shapes and colors. Left: The star-shaped red nodes represent common friends between ego users. The thickness of gray nodes represents the corresponding user's friends. Right: The nodes are colored according to the community they belong to, calculated by the modularity-based algorithm of the Louvain method. We can see several users have plural communities. The link density between communities tended to be higher within a user than between the users.

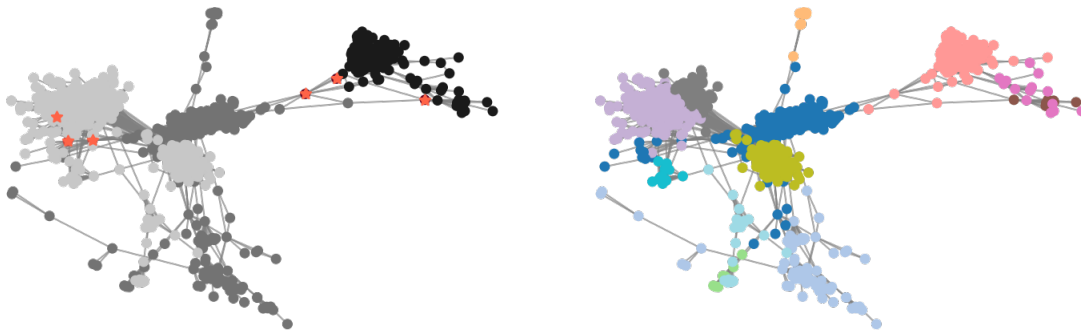


図 A.14: Facebook three ego users' network with 12 communities. Symbols and colors are as in Figure 8.

	All (10 users)				G026 (3 users)			
	P_n	P_v	PW_n	PW_v	P_n	P_v	PW_n	PW_v
0	2359	0.719	804	3.144	650	0.587	1307	2.565
1	641	0.691	2902	2.553	635	0.581	243	2.072
2	2007	0.667	3872	2.54	1141	0.579	296	2.001
3	2508	0.667	541	2.52	1249	0.568	295	1.907
4	2936	0.667	3562	2.298	1240	0.559	44	1.9
5	3244	0.667	891	2.269	1089	0.555	264	1.818
6	3583	0.661	921	2.198	935	0.551	1272	1.797
7	3081	0.658	1880	2.198	829	0.542	1195	1.797
8	3609	0.656	2130	2.198	1132	0.541	236	1.657
9	1050	0.643	776	2.19	44	0.539	225	1.556
10	2627	0.64	1859	2.186	1196	0.539	214	1.554
11	3833	0.64	145	2.129	1129	0.538	281	1.554
12	1897	0.639	377	2.088	940	0.536	533	1.546
13	3981	0.628	61	2.041	845	0.535	83	1.533
14	783	0.625	3081	1.865	1092	0.527	224	1.441
15	937	0.625	116	1.791	1073	0.526	521	1.374
16	1048	0.625	2394	1.772	549	0.524	310	1.373
17	1551	0.625	2814	1.772	1248	0.518	1083	1.317
18	2390	0.625	3469	1.693	602	0.516	249	1.248
19	3610	0.625	2038	1.649	814	0.514	374	1.192

表 A.8: **Node values in the Facebook networks with P and PW.** The values of the top 20 nodes in the two kinds of Facebook networks are shown in the order of the node ranking calculated by the two indexes of P_v and PW_v with their node IDs of P_n and PW_n .

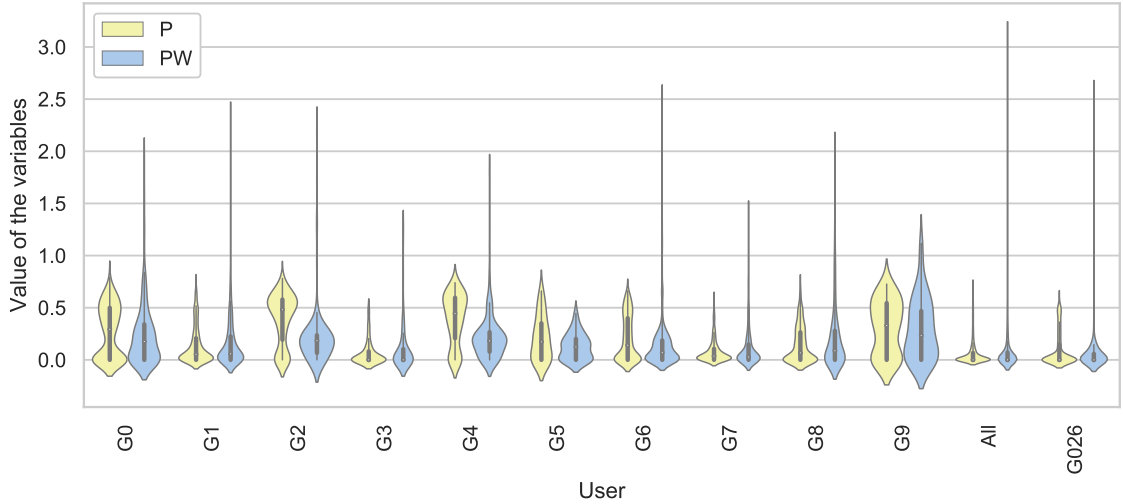


图 A.15: **Comparison of the distributions of P and PW in Facebook networks.** The violin plots represent the distributions of P and PW for each of 10 Facebook users (G0-G9), and two combined networks of all users and three users (G0, G2, and G6). The values of the X-axis and Y-axis are as in Figure 1.

ID	Z	Katz	Bet	P	PW	Burt
0	0.905	0.208	0.007	0.000	-0.000	0.416
1	-0.302	0.227	0.046	0.375	0.224	0.310
2	0.905	0.246	0.059	0.346	0.207	0.293
3	-1.508	0.226	0.196	0.531	0.632	0.258
4	-1.508	0.204	0.039	0.408	0.244	0.295
5	0.905	0.208	0.007	0.000	-0.000	0.416
6	0.905	0.208	0.007	0.000	-0.000	0.416
7	-0.302	0.189	0.001	0.000	-0.000	0.436
8	0.258	0.207	0.025	0.245	0.146	0.324
9	0.258	0.229	0.091	0.406	0.324	0.283
10	-0.775	0.211	0.037	0.408	0.244	0.287
11	0.258	0.184	0.004	0.000	-0.000	0.405
12	1.291	0.258	0.178	0.420	0.441	0.249
13	1.291	0.204	0.006	0.000	-0.000	0.390
14	-0.775	0.213	0.044	0.408	0.244	0.275
15	-1.807	0.163	0.022	0.320	0.191	0.360
16	-0.626	0.204	0.087	0.449	0.595	0.328
17	0.209	0.198	0.037	0.245	0.257	0.389
18	1.043	0.185	0.031	0.000	-0.000	0.425
19	1.043	0.185	0.031	0.000	-0.000	0.425
20	0.209	0.216	0.157	0.406	0.538	0.308
21	0.209	0.198	0.037	0.245	0.257	0.389
22	0.209	0.166	0.020	0.000	-0.000	0.446
23	-2.294	0.111	0.000	0.000	-0.000	0.565

表 A.9: **Node values in the RC networks with six indexes.** The values of all 24 nodes in the RC network are represented in the order of node ID calculated by the six indexes of Z, Kats, Bet, P, PW, and Burt. For the network visualization of node ranking, we used the order of inverse Burt, according to the definition of Burt ' s constraint and structural holes.

Rank	Z_n	Z_v	$Katz_n$	$Katz_v$	Bet_n	Bet_v	P_n	P_v	PW_n	PW_v	$Burt_n$	$Burt_v$
0	100	3.155	100	0.371	97	0.116	2	0.375	30	0.804	23	0.555
1	98	2.633	98	0.325	99	0.111	3	0.375	39	0.804	33	0.537
2	99	2.591	94	0.278	98	0.107	6	0.375	34	0.716	31	0.516
3	97	2.091	87	0.267	100	0.096	12	0.375	50	0.615	44	0.506
4	96	1.924	90	0.266	96	0.082	13	0.375	60	0.615	11	0.489
5	88	1.890	81	0.226	85	0.078	15	0.375	6	0.559	1	0.478
6	95	1.757	77	0.190	93	0.074	16	0.375	13	0.559	9	0.469
7	85	1.638	71	0.184	89	0.072	18	0.375	62	0.548	19	0.460
8	86	1.638	78	0.165	72	0.056	20	0.375	55	0.548	45	0.418
9	68	1.633	49	0.160	82	0.048	22	0.375	37	0.477	50	0.400
10	69	1.633	53	0.149	88	0.046	72	0.340	51	0.460	60	0.400
11	90	1.589	58	0.137	95	0.044	34	0.320	68	0.442	30	0.395
12	94	1.589	57	0.134	92	0.043	36	0.320	69	0.442	4	0.368
13	93	1.257	56	0.129	69	0.040	24	0.320	72	0.438	7	0.364
14	91	1.257	54	0.118	86	0.038	25	0.320	41	0.414	14	0.360
15	92	1.257	59	0.110	94	0.038	30	0.320	47	0.414	21	0.346
16	89	1.090	32	0.108	30	0.037	32	0.320	48	0.414	34	0.340
17	87	1.067	26	0.108	50	0.034	37	0.320	85	0.378	27	0.340
18	44	1.029	51	0.105	62	0.034	39	0.320	86	0.378	22	0.339
19	50	1.029	27	0.102	91	0.033	40	0.320	78	0.340	38	0.338

表 A.10: **Node values in the LFR networks with six indexes.** The values of the top 20 nodes in the LFR network are shown in the order of the node ranking calculated by the six indexes of Z_v , $Katz_v$, Bet_v , P_v , PW_v , and $Burt_v$ with their node IDs of Z_n , $Katz_n$, Bet_n , P_n , PW_n , and $Burt_n$.

Rank	Z_n	Z_v	$Katz_n$	$Katz_v$	Bet_n	Bet_v	P_n	P_v	PW_n	PW_v	$Burt_n$	$Burt_v$
0	23	4.704	16	0.302	25	0.131	295	0.708	52	1.193	177	0.200
1	59	4.241	171	0.231	23	0.120	185	0.702	35	1.141	203	0.200
2	29	3.537	24	0.180	35	0.072	131	0.700	197	0.914	294	0.200
3	22	3.289	25	0.172	22	0.058	134	0.700	131	0.795	30	0.167
4	93	3.289	21	0.169	36	0.051	133	0.698	171	0.775	96	0.167
5	5	3.216	17	0.168	93	0.050	161	0.694	133	0.771	109	0.167
6	4	3.131	23	0.164	16	0.041	36	0.694	239	0.737	135	0.167
7	73	3.032	26	0.147	39	0.041	239	0.691	269	0.718	184	0.167
8	28	3.009	304	0.143	51	0.037	105	0.685	51	0.705	244	0.165
9	2	2.960	204	0.140	171	0.031	130	0.678	248	0.693	151	0.164
10	7	2.960	22	0.139	21	0.029	52	0.672	316	0.675	235	0.145
11	16	2.631	36	0.138	67	0.027	277	0.667	169	0.667	34	0.143
12	14	2.534	211	0.130	1	0.027	171	0.665	179	0.657	128	0.143
13	11	2.449	48	0.130	209	0.027	149	0.660	295	0.650	232	0.139
14	102	2.367	43	0.128	13	0.025	201	0.656	79	0.647	237	0.132
15	61	2.305	179	0.127	4	0.022	102	0.653	185	0.640	169	0.128
16	3	2.193	147	0.126	52	0.022	21	0.653	161	0.631	125	0.128
17	25	2.132	18	0.126	2	0.019	35	0.643	30	0.624	145	0.125
18	98	2.049	45	0.126	84	0.019	249	0.640	39	0.620	162	0.125
19	1	2.022	46	0.124	5	0.018	126	0.640	25	0.612	180	0.125

表 A.11: **Node values in the Tohoku networks with six indexes.** The values of the top 20 nodes in the Tohoku network are shown in the order of the node ranking calculated by the six indexes of Z_v , $Katz_v$, Bet_v , P_v , PW_v , and $Burt_v$ with their node IDs of Z_n , $Katz_n$, Bet_n , P_n , PW_n , and $Burt_n$.

A.3 第5章の補足資料

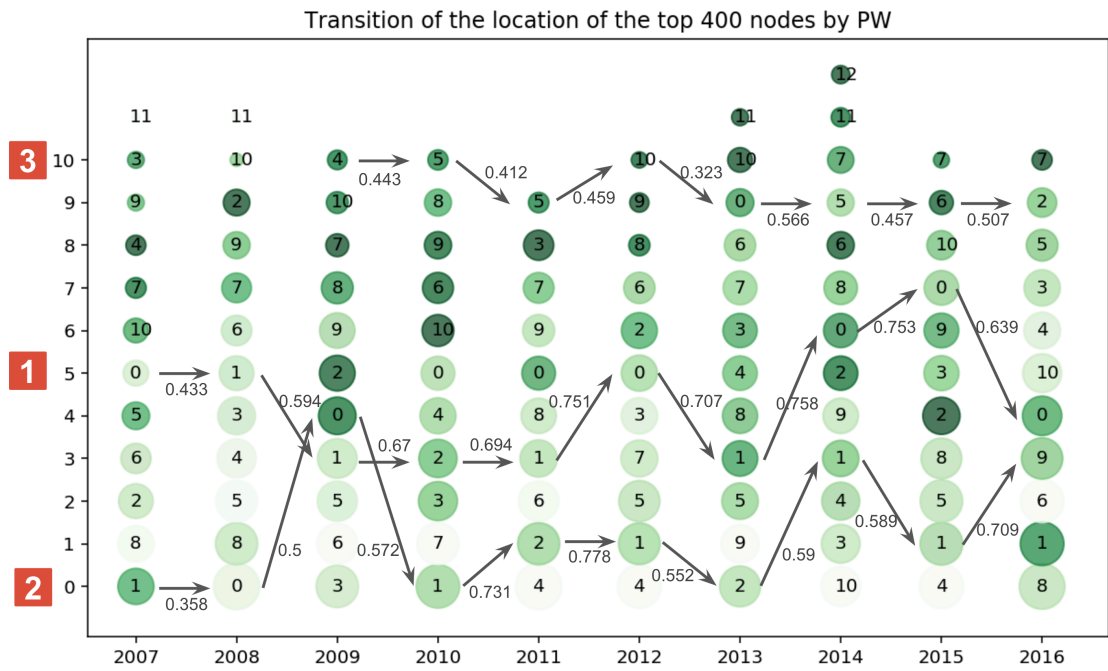


図 A.16: PW 値トップ 400 企業の分布とクラスターパス

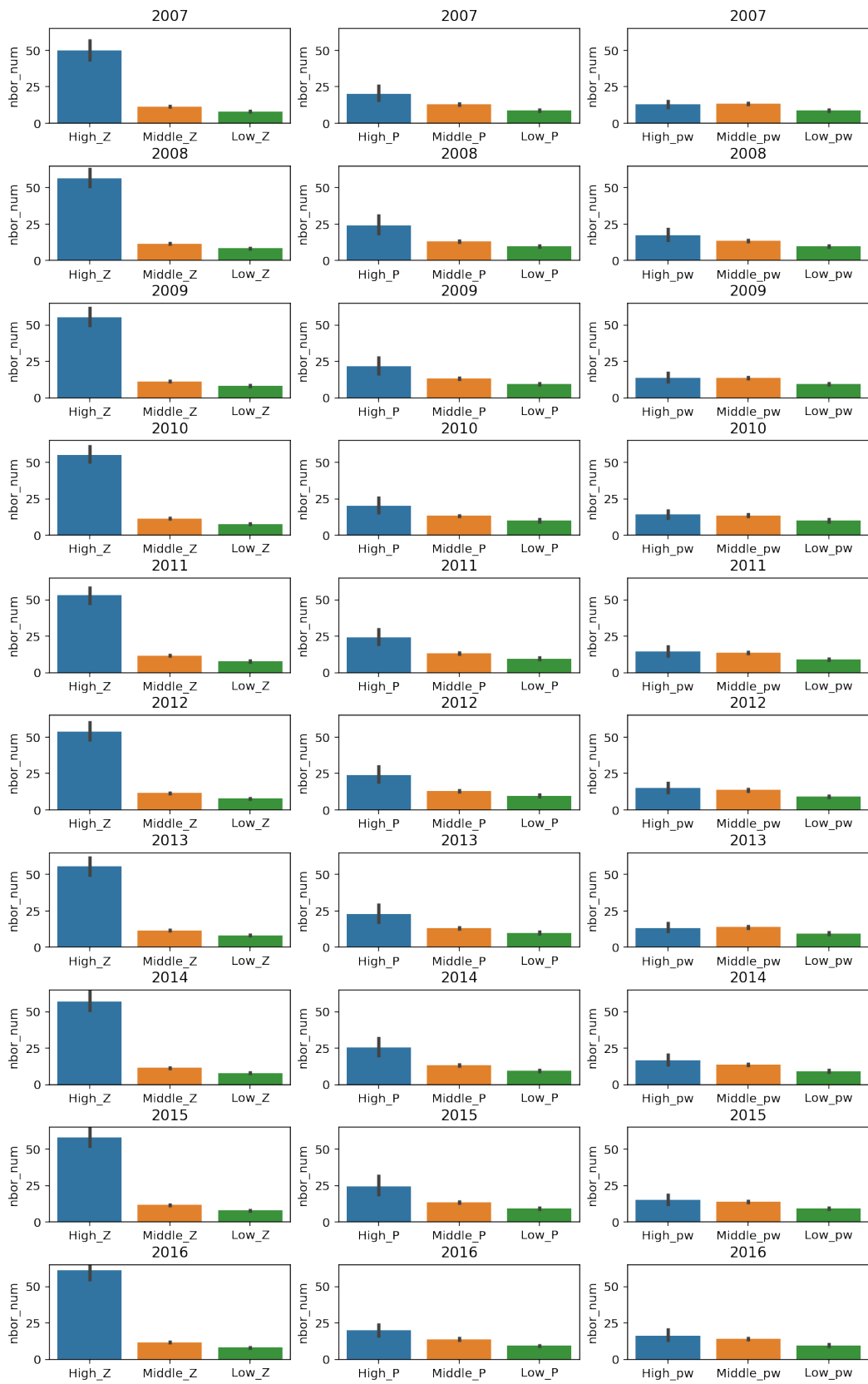


図 A.17: 各年における東北地方の企業の指標ランク別取引数

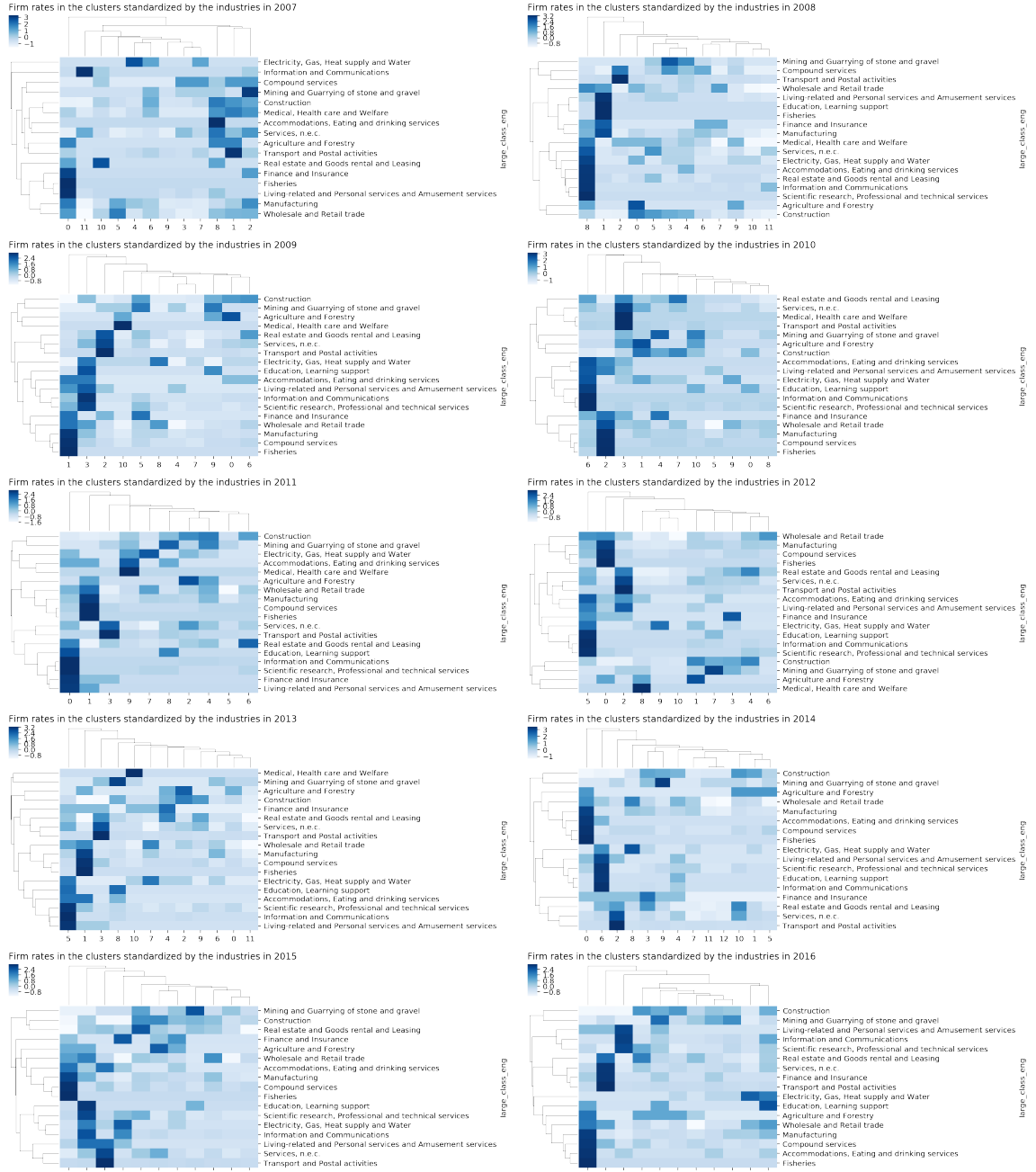


図 A.18: 産業別・クラスター別の企業割合（産業毎に標準化）

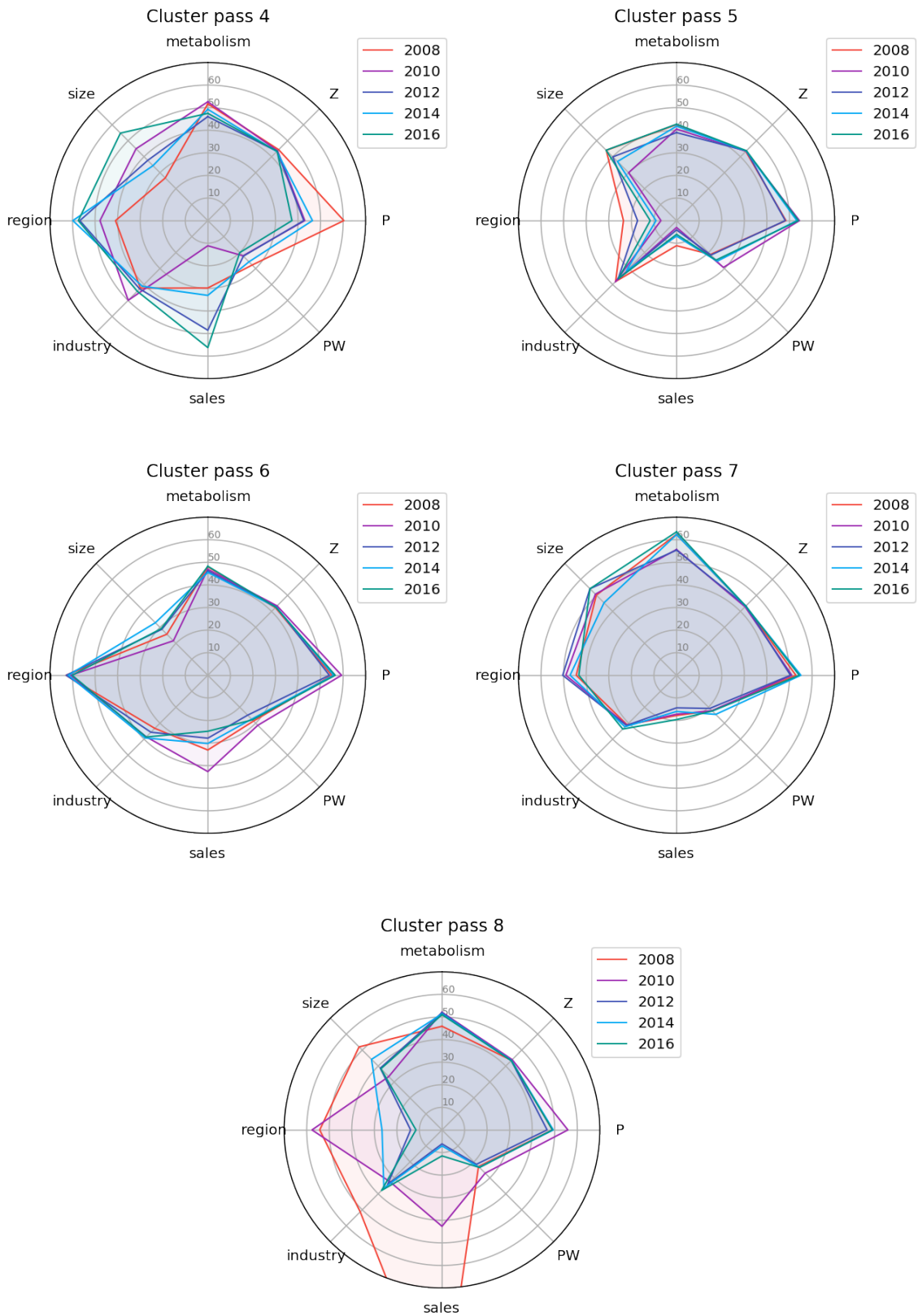


図 A.20: クラスターパスの特徴推移 (その他)

表 A.13: クラスターパス 1~6 の特徴量 (正規化・スケーリング済)

		Metabolism	Z	P	PW	Sales	Industry	Region	Size
Cluster pass 1	2008	36.63	43.85	18.85	12.5	53.87	35.9	55.65	34.9
	2010	36.18	43.74	17.69	10.7	45.04	34.79	56.93	42.9
	2012	34.99	43.68	18.01	11.4	51.82	35.2	57.69	36.9
	2014	34.52	43.58	21.71	14.21	63.69	38.02	56.65	33.7
	2016	38.02	43.99	22.25	13.44	57.52	38.95	57.86	44.2
Cluster pass 2	2008	51.53	44.39	43.11	19.66	7.54	29.18	25.78	62.3
	2010	48.84	44.73	40.36	19.71	1.78	32.24	9.26	55.6
	2012	50.56	44.7	42.6	21.02	2.43	32.3	4.57	51.1
	2014	50.32	44.5	45.48	24.14	2.8	28.77	6.69	38.7
	2016	53.61	44.33	44.3	23.25	2.79	30.29	3.21	50.0
Cluster pass 3	2010	54.13	42.05	62.66	34.09	6.56	38.12	41.9	11.7
	2012	57.88	42.92	68.65	43.44	5.91	28.92	21.1	7.2
	2014	64.59	42.22	61.8	30.35	5.54	30.83	31.06	20.8
	2016	62.41	42.69	64.71	30.27	7.37	37.37	31.7	24.6
Cluster pass 4	2008	51.71	44.43	60.33	27.82	29.85	42.29	40.93	26.7
	2010	52.63	43.74	43.01	22.06	11.13	49.93	47.82	45.0
	2012	46.05	43.39	42.54	22.27	48.55	42.74	56.76	37.7
	2014	49.29	43.28	46.41	25.81	33.11	41.12	59.85	34.4
	2016	47.5	43.45	37.32	20.04	56.24	44.24	57.43	54.8
Cluster pass 5	2008	42.54	43.34	48.44	21.21	11.1	38.16	23.5	44.0
	2010	40.51	43.49	54.42	29.33	3.03	38.3	6.91	30.0
	2012	38.97	43.7	48.14	21.56	4.12	36.62	17.27	39.9
	2014	41.8	43.63	53.27	25.44	6.89	31.44	9.09	37.0
	2016	42.69	43.76	53.94	24.72	6.08	36.33	11.65	44.1
Cluster pass 6	2008	48.05	42.35	54.96	28.66	33.1	33.24	62.84	25.7
	2010	46.93	43.46	59.31	31.35	42.65	38.58	60.79	21.6
	2012	46.22	42.89	53.89	24.99	27.87	35.68	62.8	28.8
	2014	45.53	42.72	55.79	27.48	30.22	39.28	61.88	32.9
	2016	48.27	42.64	56.51	27.93	24.81	38.67	60.16	29.2
Cluster pass 7	2008	62.4	43.19	52.84	22.24	17.82	30.98	44.42	50.2
	2010	55.71	42.89	51.04	22.18	17.26	31.17	48.99	50.8
	2012	55.4	43.42	50.43	20.77	14.42	31.4	50.4	54.2
	2014	62.28	43.26	55.05	24.51	16.04	32.29	47.13	45.4
	2016	63.5	43.33	54.34	22.4	19.63	33.68	43.24	54.2
Cluster pass 8	2008	45.91	43.57	48.93	22.9	127.95	51.29	54.23	52.0
	2010	52.14	44.01	55.84	27.08	42.68	32.53	57.65	33.4
	2012	51.93	43.43	46.62	21.54	6.18	34.09	13.92	38.7
	2014	51.38	43.54	49.36	23.37	7.08	36.45	26.61	44.2
	2016	51.03	43.42	48.99	23.53	11.5	37.92	11.54	38.3

表 A.14: クラスタースパス 4 の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	2	6	2	3	3	2	3	2	2	1
サービス業その他	13	5	26	33	17	36	27	30	39	77
卸売業, 小売業	70	59	62	87	36	89	52	62	72	144
建設業	182	120	54	109	17	7	10	10	5	5
製造業	48	65	33	29	21	20	31	27	32	31
複合サービス事業	2	4	2	1	2	2	5	2	4	6
運輸業, 郵便業	8	6	167	164	163	196	200	194	231	250
宮城県	14	22	79	92	64	116	84	76	94	144.0
山形県	2	3	37	41	29	66	40	37	39	65.0
岩手県	6	19	31	9	13	25	41	41	46	51.0
東京都	13	21	21	14	17	21	23	27	30	29.0
福島県	7	5	87	216	52	73	63	75	86	142.0
秋田県	164	32	28	23	27	20	22	22	34	34.0
青森県	137	161	74	43	56	43	61	57	67	70.0

表 A.15: クラスタースパス 5 の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	3	4	9	10	8	7	8	9	9	10
サービス業その他	7	9	6	7	8	9	6	5	5	9
卸売業, 小売業	14	78	57	52	59	66	53	47	43	61
建設業	45	268	251	183	216	249	243	260	250	295
製造業	9	47	36	32	46	47	35	36	33	39
運輸業, 郵便業	1	20	3	12	20	12	7	7	8	9
鉱業, 採石業, 砂利採取業	1	5	5	4	5	5	6	6	5	5
宮城県	53	44	5	6	9	22	9	11	8	15.0
山形県	2	2	-	-	-	1	-	-	-	-
岩手県	5	349	338	282	329	347	342	346	336	406.0
東京都	8	17	6	7	8	10	6	9	8	9.0
福島県	6	11	1	-	-	4	-	-	-	1.0
秋田県	3	3	-	-	1	2	-	-	-	1.0
青森県	5	10	9	-	11	9	3	1	1	3.0

表 A.16: クラスタースパス 6 の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	5	7	8	9	7	6	7	8	8	5
不動産業, 物品賃貸業	1	1	6	6	5	10	5	8	12	4
卸売業, 小売業	106	146	162	111	138	154	156	150	161	131
建設業	61	53	68	52	70	66	108	93	226	81
製造業	50	51	53	38	52	53	61	64	72	64
宮城県	48	58	65	67	70	66	68	81	69	92.0
山形県	26	35	39	26	36	39	33	39	26	30.0
岩手県	26	28	33	33	42	35	48	52	54	36.0
東京都	20	22	23	18	22	23	22	24	26	26.0
福島県	44	45	69	27	49	55	50	55	40	49.0
秋田県	21	27	21	12	17	21	26	18	21	16.0
青森県	25	28	29	20	27	30	75	40	234	26.0

表 A.17: クラスターパス7の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	8	5	6	7	6	4	9	10	4	6
サービス業その他	13	22	9	11	8	9	13	18	14	21
不動産業, 物品質貸業	1	3	14	17	15	13	11	18	21	21
卸売業, 小売業	60	82	88	99	100	108	71	68	103	89
建設業	260	351	321	338	298	354	295	312	360	364
製造業	20	28	33	38	39	52	22	30	49	31
複合サービス事業	1	2	3	2	2	2	3	2	3	4
宮城県	184	266	227	220	206	207	177	195	233	225.0
山形県	9	19	16	24	10	13	11	12	12	9.0
岩手県	21	27	21	30	22	20	36	16	35	27.0
東京都	31	34	39	39	44	51	39	40	46	35.0
福島県	98	112	114	141	88	145	93	134	160	206.0
秋田県	13	13	10	18	20	20	13	13	18	11.0
青森県	14	20	39	26	64	81	42	37	42	19.0

表 A.18: クラスターパス8の産業別・地域別企業数

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
継承クラスター ID	10	8	3	0	4	3	4	3	3	3
サービス業その他	4	33	12	7	9	8	8	8	8	11
不動産業, 物品質貸業	2	10	8	3	9	10	15	18	7	7
卸売業, 小売業	45	163	151	81	122	55	52	68	63	69
建設業	102	171	218	202	392	262	234	278	236	238
製造業	16	85	76	37	81	35	29	55	33	36
複合サービス事業	1	1	4	1	2	2	2	2	2	2
運輸業, 郵便業	1	5	8	1	3	8	5	6	4	10
宮城県	51	222	226	117	88	2	8	9	8	5.0
山形県	5	49	33	18	13	2	3	2	2	4.0
岩手県	8	59	35	39	32	3	5	3	2	-
東京都	22	47	55	35	36	10	10	13	8	12.0
福島県	12	57	95	44	19	-	3	4	-	2.0
秋田県	54	38	32	25	358	348	310	322	337	355.0
青森県	20	32	39	44	62	21	9	85	3	2.0

A.4 ソースコード

第4章で用いたノード指標 Z,P,PW 値を計算し、ネットワーク上にノード順位と共に描画するプログラムを、ソースコード A.1 に示す。

(1) 隣接クラスター行列

エッジリストから隣接行列を取得。隣接行列の配列がグラフのノード順と一致するよう並び替える。グラフの全ノードについて、所属クラスター ID リストを取得所属クラスター ID リストと隣接行列の内積を取ることで、要素に隣接ノードのクラスター ID が入った隣接クラスター行列を作成

(2) 非所属隣接クラスター行列

各クラスターのノード ID を取得隣接クラスター行列のうち、各ノード毎に所属クラスター ID は 0 に置換し、所属クラスターの外にある隣接ノードのクラスター ID のみを残す。その際、隣接クラスター行列の要素を、各クラスターのノード ID で指定する作業を繰り返すことで、全てのクラスターについて同様に、所属クラスターの外にある隣接ノード（非所属隣接ノード）のクラスター ID のみを抽出する。得られたクラスター ID 行列を連結し、グラフのノード順に並び替え、非所属隣接クラスター行列とする。

(3) 非所属隣接クラスター関連度行列

全てのクラスターペアについて、クラスター関連度の式に基づいてクラスター関連度辞書を取得。非所属隣接クラスター行列のクラスター ID をクラスター関連度に置き換えた、非所属隣接クラスター関連度行列を作成する。その際、各ノードについて、所属クラスターと非所属隣接ノードの所属クラスターとの関連度を求める作業を、ノードの所属クラスター毎に繰り返すことで、全てのクラスターについて同様に、所属クラスターの外にある隣接ノードのクラスター関連度を抽出する。得られたクラスター関連度行列を連結し、グラフのノード順に並び替え、非所属隣接クラスター関連度行列とする。

(4) 非所属隣接クラスター非関連度

各ノードのクラスター外次数を求める。非所属隣接クラスター関連度行列から、ノード毎に非所属隣接クラスター関連度の合計を求め、クラスター外次数で割ることで、クラスター関連度の平均値を取得する。この際、ノードのクラスター外次数が 0 である場合に対応するため、分母に $1e-7$ を加え inf 回避（分子が 0 なので結果的に 0 になる）。平均クラスター関連度の対数を負にした値を非関連度とする。平均クラスター関連度が 0 の場合は、0 を 1 に置き換えた値を非関連度とする。

(5) 重み付き P 値

各ノードの次数を取得する隣接クラスター行列を用いて、ノード毎に各クラスターに属する隣接ノード数をノードの次数で割った値の二乗の総和を求めることで P 値を取得する。ノード毎の P 値に非所属隣接クラスター非関連度を掛けることで、重み付き P 値を取得する

ソースコード A.1: calc-pw

```
1 # coding: utf-8
2 '''
3 [メインプログラムとして実行]
4
5 1. 計算してpkl を保存
6 $ python calc_weighted_p.py calc input.edges output.pkl
7
8 2. pkl からグラフを描画
9 $ python calc_weighted_p.py draw output.pkl output.eps
10
11 3. 計算&描画
12 $ python calc_weighted_p.py calc-draw input.edges output.eps
13
14 [ライブラリとして使用]
15 以下の関数を使用可能
16
17 calc_weighted_p(G)
18     [引数]
19     G: 計算するネットワーク (networkx graph object)
20     [返回值]
21     WeightedP object
22     -----
23     class WeightedPVals:
24         def __init__(self):
25             self.z = 0
26             self.p = 0
27             self.zp = 0
28             self.p_weight = 0
29             self.zp_weight = 0
30             self.between = 0
31             self.katz = 0
32             self.G = 0
33             self.burt = 0
34             return
35     -----
36
37 draw_weighted_p(weighted_p, savepath)
38     [引数]
39     weighted_p: WeightedP object
40     savepath: グラフの保存先
41
42 '''
```

```

43
44 import sys
45 import itertools
46 import numpy as np
47 import pandas as pd
48 import networkx as nx
49 import community
50 import pickle
51 import matplotlib.pyplot as plt
52
53
54 class WeightedPVals:
55     def __init__(self):
56         self.z = 0
57         self.p = 0
58         self.zp = 0
59         self.p_weight = 0
60         self.zp_weight = 0
61         self.between = 0
62         self.katz = 0
63         self.G = 0
64         self.burt = 0
65         return
66
67
68 # ---- Private functions ----
69
70 def _crjc(G, partition=None):
71     def predict(Cu,Cv,nbor_dic):
72         cn = nbor_dic[Cu] & nbor_dic[Cv]
73         nbor_uv = nbor_dic[Cu] | nbor_dic[Cv]
74         jc = len(cn)/len(nbor_uv)
75         return jc, cn
76     nbor_dic = {}
77     for c in set(partition.values()):
78         nodes = [x[0] for x in partition.items() if x[1]==c]
79         nbors = [set(np.array([x for x in G.edges if x[0]==u or x[1]==u
80             ]).flatten()) for u in nodes]
81         unique_nbors = []
82         for x in nbors:
83             unique_nbors.extend(list(x))
84         nbor_dic[c] = set(unique_nbors)
85     combi = list(itertools.combinations(set(partition.values()), 2))
86     dic = {}
87     for u,v in combi:
88         dic[(u+1,v+1)]=predict(u,v,nbor_dic)
89     return dic
90
91

```

```

92 def _make_one_hot(G):
93     edge = pd.DataFrame(list(G.edges))
94     edge_r = edge[[1,0]]
95     edge_r.columns=[0,1]
96     edge2 = pd.concat([edge, edge_r])
97     cross = pd.crosstab(edge2[0], edge2[1])
98     cross = cross.reindex(G.nodes, axis=0)
99     cross = cross.reindex(G.nodes, axis=1)
100    return cross
101
102
103
104 def _extract_outers(G, partition):
105     cross = _make_one_hot(G)
106     Ko = cross.sum(axis=1)
107     cii = np.array(list(partition.values()))
108     cii +=1
109     cii_diag = pd.DataFrame(np.diag(cii), index=cross.index, columns=
        cross.columns)
110     Gc = cross @ cii_diag
111     c1_nodes = [x[0] for x in partition.items() if x[1]==0]
112     c1_outer = Gc.loc[c1_nodes]
113     outer = c1_outer.where(c1_outer!=1, 0)
114     for c in range(1,len(set(cii))):
115         c_nodes = [x[0] for x in partition.items() if x[1]==c]
116         c_outer = Gc.loc[c_nodes]
117         c_outer = c_outer.where(c_outer!=c+1, 0)
118         outer = pd.concat([outer, c_outer])
119     outer = outer.reindex(G.nodes, axis=0)
120     return cii, outer
121
122
123 def _cls_sim_node(G, partition):
124     jac = _crjc(G, partition)
125     cii, outer = _extract_outers(G, partition)
126
127     c1_nodes = [x[0] for x in partition.items() if x[1]==0]
128     c1_outer = outer.loc[c1_nodes]
129     sim1 = np.where(c1_outer==2, jac[1,2][0],0)
130     for x in range(3,len(set(cii))+1):
131         sim1_ = np.where(c1_outer==x, jac[1,x][0],0)
132         sim1 = sim1 + sim1_
133     sim1 = pd.DataFrame(sim1, index=c1_nodes, columns=G.nodes)
134
135     cii_set = set(cii)
136     for c in range(2,len(cii_set)+1):
137         combi = [x for x in itertools.combinations(cii_set,2) if x[0]==c
            or x[1]==c]
138         c_nodes = [x[0] for x in partition.items() if x[1]==c-1]
139         c_outer = outer.loc[c_nodes]

```

```

140     sim2 = np.where(c_outer==1, jac[combi[0]][0], 0)
141     for i in combi[1:]:
142         if c > i[0]:
143             sim2_ = np.where(c_outer==i[0], jac[i][0], 0)
144         if c <= i[0]:
145             sim2_ = np.where(c_outer==i[1], jac[i][0], 0)
146         sim2 = sim2 + sim2_
147     sim2 = pd.DataFrame(sim2, index=c_nodes, columns=G.nodes)
148     sim1 = pd.concat([sim1, sim2])
149     sim1 = sim1.reindex(G.nodes, axis=0)
150     return sim1
151
152
153 def _cls_sim_mean_val(G, partition):
154     cii, outer = _extract_outers(G, partition)
155     sim_cls = _cls_sim_node(G, partition)
156
157     outer_one = np.where(outer!=0,1,0)
158     outer_count = np.sum(outer_one, axis=1)
159     sim_sum = np.sum(sim_cls, axis=1)
160     sim_mean = sim_sum/(outer_count + 1e-7)
161     sim_mean = np.where(sim_mean==0,1,sim_mean)
162     sim_mean = -np.log(sim_mean)
163     return sim_mean
164
165
166
167 def _calc_zp_weight(G, partition, weight):
168     cross = _make_one_hot(G)
169     W = cross.values
170     Ko = np.sum(W, axis=1)
171     cii = np.array(list(partition.values()))
172     cii +=1
173
174     Gc = np.dot(W, np.diag(cii))
175     Z = np.zeros(len(W))
176     Kc2 = np.zeros(len(W))
177     for i in range(1, int(np.max(cii)+1)):
178         select = W * (Gc == i)
179         Koi = np.sum(select[cii == i], axis=1)
180         Z[cii == i] = (Koi - np.mean(Koi)) / (np.std(Koi)+1e-7)
181         Kc2 += (np.sum(W * (Gc == i), axis=1)/Ko)**2
182
183     P = np.ones(len(W)) - Kc2
184     zp = Z*P
185     p_weight = P*weight
186     zp_weight = Z*p_weight
187
188     Z = dict(zip(G.nodes, Z))
189     P = dict(zip(G.nodes, P))

```



```

190     zp = dict(zip(G.nodes, zp))
191     p_weight = dict(zip(G.nodes, p_weight))
192     zp_weight = dict(zip(G.nodes, zp_weight))
193     return Z, P, zp, p_weight, zp_weight
194
195
196
197 # ---- Public functions ----
198
199 # -----
200 # Calculation of weighted P values from the graph
201 # G: networkx graph
202 # -----
203
204 def calc_weighted_p(G):
205     weighted_p = WeightedPVals()
206     partition = community.best_partition(G)
207     weighted_p.between = nx.betweenness_centrality(G)
208     weighted_p.katz = nx.katz_centrality_numpy(G)
209     weighted_p.burt = nx.constraint(G,G.nodes)
210     sim = _cls_sim_mean_val(G, partition)
211     weighted_p.z, weighted_p.p, weighted_p.zp, weighted_p.p_weight,
        weighted_p.zp_weight = _calc_zp_weight(G, partition, sim)
212     weighted_p.G = G
213     return(weighted_p)
214
215
216 def draw_weighted_p(weighted_p, savepath):
217     G = weighted_p.G
218     vals = [weighted_p.z, weighted_p.katz, weighted_p.between, weighted_p.
        p, weighted_p.p_weight, weighted_p.burt]
219     names = ['Z', 'Katz', 'Betweenness', 'P', 'P_weight', 'iBurt']
220     pos = nx.spring_layout(G, seed=1)
221     fig = plt.figure(figsize=(16,10))
222     for i,pr in enumerate(vals):
223         ax = fig.add_subplot(3,3,i+1)
224         nx.draw_networkx_edges(G, pos, edge_color='Gray', alpha=0.7)
225         if i<5:
226             sorted_value = [k for k, v in sorted(pr.items(), key=lambda x
                : x[1], reverse=True)][:10]
227             ranking = dict([(x[1], x[0]) for x in enumerate(sorted_value
                )])
228             nx.draw_networkx_nodes(G, pos, node_size=30, node_color=list(
                pr.values()), cmap=plt.cm.Greens)
229             nx.draw_networkx_labels(G, pos, ranking)
230         elif i==5:
231             sorted_value = [k for k, v in sorted(pr.items(), key=lambda x
                : x[1], reverse=False)][:10]
232             ranking = dict([(x[1], x[0]) for x in enumerate(sorted_value
                )])

```

```

233         nx.draw_networkx_nodes(G, pos, node_size=30, node_color=[1/(v
                +1e-7) for v in pr.values()], cmap=plt.cm.Greens)
234         nx.draw_networkx_labels(G, pos, ranking)
235         ax.set_title(names[i])
236         plt.axis('off')
237     plt.show()
238     fig.savefig(savepath)
239     return
240
241
242 # -----
243 # Main routine
244 # -----
245
246 def main():
247     if ( len(sys.argv)!=4 ):
248         print("Error: Invalid argument", file=sys.stderr)
249         exit(1)
250
251     if ( sys.argv[1]=="calc" ):
252         G = nx.read_edgelist(sys.argv[2])
253         weighted_p = calc_weighted_p(G)
254         with open(sys.argv[3], 'wb') as f:
255             pickle.dump(weighted_p, f)
256
257     elif ( sys.argv[1]=="draw" ):
258         with open(sys.argv[2], 'rb') as f:
259             weighted_p = pickle.load(f)
260             draw_weighted_p(weighted_p, sys.argv[3])
261
262     elif ( sys.argv[1]=="calc-draw" ):
263         G = nx.read_edgelist(sys.argv[2])
264         weighted_p = calc_weighted_p(G)
265         draw_weighted_p(weighted_p, sys.argv[3])
266
267     exit(0)
268     return
269
270
271 if __name__ == "__main__":
272     main()

```

本論文に関連した研究業績

1. **H. Yamano**, K. Asatani, and I. Sakata, "Evaluating Nodes of Latent Mediators in Heterogeneous Communities", *Scientific Reports* 10 (2020), 8456. doi:10.1038/s41598-020-64548-6.
2. **H. Yamano** and I. Sakata, "Roles of Brokers and Clusters in the Inter-firm Network Dynamics: Evolution Map Perspective", In *Proceedings of the 2022 Portland International Conference on Management of Engineering and Technology (PICMET)*, in Portland USA (August 7-11, 2022), pp. 1-10, forthcoming.
3. **H. Yamano**, K. Asatani, and I. Sakata, "Evaluating Nodes of Latent Mediators in Heterogeneous Communities", In *Book of Abstracts of the 8th International Conference on Complex Networks and their Applications (Complex Networks 2019)*, in Lisbon Portugal (December 10-12, 2019), pp. 90-92.
4. 山野泰子, 坂田一郎, 異質なコミュニティをつなぐノードの評価指標. In *WebDB Forum 2019 論文集*, Vol. 2019, (September 8-9, 2019), pp. 1-4.
5. **H. Yamano**, I. Sakata, "Latent Pattern Extraction and Factorization of Firm Bankruptcies and Metabolism in Japan", In *Proceedings of the 2019 Portland International Conference on Management Engineering and Technology (PICMET)*, in Portland USA (August 25-29, 2019), pp. 1-12, doi:10.23919/picmet.2019.8893943.
6. **H. Yamano**, I. Sakata, "Regional Differences and Commonalities of Inter-Firm Transaction Metabolism", In *Proceedings of the 2017 International Society for Professional Innovation Management (ISPIM) Innovation Summit*, in Melbourne Australia (December 10-13, 2017), pp. 1-13.
7. **H. Yamano**, H. Sasaki, and I. Sakata, "Metabolism of inter-firm transactions in regional networks", In *Proceedings of the 2017 Portland International Conference on Management Engineering and Technology (PICMET)*, in Portland USA (July 9-13, 2017), pp. 1-8, doi:10.23919/picmet.2017.8125409.