

極大クリーク全列挙アルゴリズムを用いた企業コミュニティ理解

米 森 力† 松 永 務† 富 田 悦 次††

これまで企業評価にあたり、対象企業の取引に関する評価が主に当該取引先の範囲で行われているのに対し、当該企業の属すコミュニティから捉えた評価が望まれている状況にある。本論文では、企業コミュニティ理解が、企業間の取引情報において企業をノード、取引関係をエッジとした大規模グラフからのクリーク抽出により実現されることを示す。全列挙された極大クリークに基づき、企業コミュニティを捉え、クリークに含まれるノード数から企業の取引構造の特徴を定量化できる指標を新たに考案した。東証一部上場企業 2,903 社を対象にした実験から有用性を示す。互いに取引し合う密型と特定企業を中心に取引が行われる星型との取引構造があり、考案指標がその取引構造を表現し、企業評価に有効に機能することを明らかにした。

An analysis of enterprise communities by cliques.

CHIKARA YONEMORI†, TSUTOMU MATSUNAGA†
and ETSUJI TOMITA ††

Although assessment of enterprise with regard to trading has been examined, it is the situation that the assessment of trading partner enterprises as communities is needed. In this paper, it is shown that the enterprise communities can be treated as cliques, which are extracted from a large-scale graph where nodes and edges correspond to enterprises and trades, respectively. We devised a novel measure which quantizes a feature of trade structure using the cliques. Experimental results for 2,903 enterprises show the enterprise community type has star structure and dence structure based on the measure. Thus the measure is effective for analysing the trading structures.

1. はじめに

企業評価は、取引時の与信をはじめ、昨今の企業の吸収合併を代表例に一層重要度を増し、情報分析の対象として期待されている。ここに、企業間の関係は、経営上の意志決定に関する重要な情報源として活用される。一般に、投資先の決定や融資先の限度額の決定では、リスク把握の観点から対象となる企業の資本関係や提携関係が考慮される¹⁾がこれまでは取引先のみ把握に留まり、取引先の取引先など隣接した企業間関係を超越するような全体的な視点からは、取引関係を理解することができない状況にあった²⁾。取引先のみではなく取引構造が与信等の企業評価における意思決定に今後ますます重要となると考えられる。

本論文では、極大クリーク的全列挙によるグラフ構造の特徴の定量化を通して、企業の取引構造の把握を図る。企業をノード、企業間の取引関係をエッジとするグラフに対して、極大クリーク全列挙アルゴリズムを適用する。個々の企業におけるグラフ構造の特徴をクリークのサイズに着目した指標により算出し、企業の

取引構造が企業コミュニティの視点から定量化される。

以下、2. で企業コミュニティ抽出における企業間関係ネットワークについて述べ、3. で提案する分析方法について明らかにする。4. では、実際の企業情報を用いた実験評価を通して、提案法の有効性を示す。5. はまとめと今後の課題である。

2. 企業間関係におけるネットワーク分析

従来のネットワークに対する分析方法としてコミュニティ抽出が挙げられ³⁾、また、グラフ構造に応じた部分グラフの抽出やクリークの列挙のアプローチ^{4),5)}が知られている。

企業の活動は一般的に企業間の関係から記述される。例えば企業間の取引活動においては、何らかの材料を仕入れて、製品やサービスを生みだし、販売先に提供するという商取引が成立している。こうした活動は、企業間取引のネットワークとして、グラフで記述することができる。極大クリークはノード間が全てエッジで結ばれる完全部分グラフであり、企業間取引関係においては密な取引構造を持つ企業コミュニティに対応する。従って、個々の企業に着目したとき、当該企業がノード数の大きい極大クリークに属しているほど、密な構造を持つコミュニティに属していると思われることができ、属するクリークの大小に関する構成比を見ることができ、当該企業の持つ取引関係構造のタイプ分類に繋がる定量化が期待される。

† 株式会社 NTT データ 技術開発本部
NTT DATA Corporation, Research and Development
Headquarters.

†† 電気通信大学 電気通信学部情報通信工学科
The University of Electro-Communications, Department
of Information and Communication Engineering.

3. クリークを用いた企業コミュニティ理解

3.1 極大クリーク全列挙アルゴリズムの適用

グラフ G の部分グラフのうち、完全グラフとなっているものはクリークと呼ばれ、そのうち、真に小さいクリークをその中に含まないクリークは極大クリークと呼ばれる。一般に、極大クリーク全列挙問題は NP 困難な問題として知られる⁴⁾⁵⁾。ここでは、極大クリークを列挙する方法として極大クリーク全列挙アルゴリズム CLIQUES⁴⁾を用いる。大規模なグラフで最も高速に処理しうることが確認されており、本検討においてもふさわしいと考えた。

本節では、ノードの有限集合 V 、異なるノードの対 $(v, w) \{v \in V, w \in V, v \neq w\}$ を接続するエッジの有限集合 E より構成される無向グラフを $G = V \times E$ と表記する。 G に関して極大クリークを全列挙した結果集合を $C = \{C_1, \dots, C_m\}$ とする。ここで m は抽出された極大クリーク数である。また極大クリーク C_i のサイズ $|C_i|$ は、極大クリークに含まれるノード数とする。

3.2 擬似クリークによる企業コミュニティ抽出

前節で述べた極大クリークにおいて、構成ノードがわずかに異なる極大クリークが多数出力されることが指摘されており、こうしたわずかな違いを吸収するため、擬似クリークの列挙というアプローチ⁶⁾が知られている。擬似クリークは、極大クリーク同士の重複を許すように極大クリークの定義を拡大することで、極大クリーク同士を統合し、より解釈のしやすいクリークを構成するものである。

ある分析対象のグラフ G より得られた極大クリークの集合 C' の重複度を次式

$$\text{overlap}(C') = \min_{C'_i \in C'} \left(\frac{|\bigcap_{C'_h \in C'} C'_h|}{|C'_i|} \right) \quad (1)$$

で定義する。ここで \cap はクリークの積集合を表し、重複度は定義より、 $0 \leq \text{overlap}(C') \leq 1$ の値をとる。与えられた極大クリーク C' に関する擬似クリーク集合 $PC(C')$ は、極大クリークの組み合わせとして、次式

$$PC(C') = \bigcup_{C'_h \in C'} C'_h, \text{overlap} \left(\bigcup_{C'_h \in C'} C'_h \right) > \alpha \quad (2)$$

より定義される。ここで \cup はクリーク之和集合を表し、 α は重複度を制御する閾値であり、この閾値を超えるように擬似クリークが構成される。ここでは、得られた擬似クリークを企業コミュニティとみることとする。

3.3 結束度に基づく企業コミュニティ理解

本節では、企業コミュニティの抽出結果を用いて、企業の持つ取引構造をタイプ分類するための指標として、結束度を定義する。

結束度の算出ではまず、グラフ G に含まれるノード v に関して、極大クリーク C'_i のサイズ $|C'_i|$ に含まれる頻度 $f_{(v, |C'_i|)}$ を要素を持つ行列を次式

$$F = f_{(v, |C'_i|)}, v \in V, |C'_i| = 1, \dots, \max(|C'_i|) \quad (3)$$

とし、あるノード v の結束度 u_v は、次式

$$u_v = \frac{\sum_k \{f_{(v, k)}\} - f_{(v, 2)}}{\sum_k \{f_{(v, k)}\}}, k \leq 2 \quad (4)$$

で算出され、この値により当該ノードに対して、接続関係が相互に密接な密型か単純な星型かのタイプ分類が可能となる。

4. 実験

4.1 実験に用いたデータ

本実験では、実際の企業データとして、国内の企業の2003年度の仕入並びに販売の取引情報(日経 NEEDS)を用いる。実験は、上場企業を対象とし、仕入先または販売先の取引関係が連結する2,903社を用いた。

次数の高い企業例を表1に挙げる。表には各企業の属する業種を示した。表1からは取引先の多い企業として、物流関係を持つ商社や系列を持つ自動車業が上位に挙げられることが示されている。

表1 高い次数を持つ企業(上位20社)の業種

順位	次数	業種
1	288	総合商社
2	234	総合商社
3	154	総合商社
4	124	総合商社
5	118	自動車
6	107	通信機(含通信機部品)
7	102	総合商社
8	91	自動車
9	76	家庭電器(含音響機器)
10	74	総合商社
11	72	印刷
12	62	自動車関連
13	62	機械金属商社
14	59	大手建設
15	57	自動車
16	54	印刷
17	54	総合商社
18	51	電気機器
19	48	家庭電器(含音響機器)
20	48	化学商社

4.2 クリークを用いたコミュニティ抽出

4.2.1 極大クリークの抽出

対象データから得られた極大クリークのうち、サイズ上位の極大クリークに含まれる企業の業種を表2に示す。得られた極大クリークの数は6,302個であり、サイズの最大値は5が得られた。表2より、サイズが大きいクリークに含まれる企業の特徴として自動車業および総合商社が含まれていることがわかる。表1との対比により、高い次数を持つ総合商社よりも自動車業がサイズの大きい極大クリークとして挙げられる傾向が示されている。

4.2.2 擬似クリークによるコミュニティ抽出

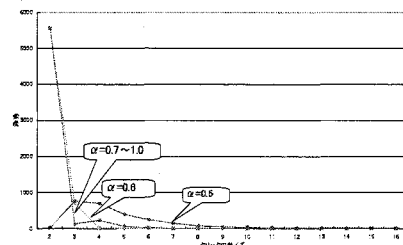


図1 クリークのサイズに対する擬似クリークの頻度

表 2 サイズの降順に並べた極大クリーク (上位 20 個) と属する企業の業種

順位	クリークのサイズ	業種名 1	業種名 2	業種名 3	業種名 4	業種名 5
1	5	内燃機	自動車	自動車関連	建材商社	自動車部品
2	5	内燃機	自動車	自動車関連	建材商社	自動車部品
3	4	総合商社	大手化学	農薬	化工機械	
4	4	総合商社	塩業・ソーダ	塗料	信化学	
5	4	自動車部品	自動車	車体・他	建材商社	
6	4	自動車部品	自動車	車体・他	建材商社	
7	4	自動車部品	自動車	自動車関連	建材商社	
8	4	大手化学	総合商社	化学商社	合成樹脂	
9	4	大手化学	総合商社	化学商社	石油精製及び販売	
10	4	他化学	自動車	自動車関連	自動車	
11	4	他ゴム製品	自動車	自動車	車体・他	
12	4	内燃機	自動車	自動車関連	自動車部品	
13	4	内燃機	自動車	自動車関連	電機関連商社	
14	4	内燃機	自動車	自動車	建材商社	
15	4	内燃機	自動車	自動車部品	建材商社	
16	4	内燃機	自動車	自動車部品	自動車部品	
17	4	自動車	自動車	自動車	自動車部品	
18	4	輪受	自動車	建材商社	自動車部品	
19	4	通信機 (含通信機部品)	電子部品	電子部品	他商社	
20	4	家庭電器 (含音響機器)	自動車	車体・他	自動車部品	

表 3 サイズの降順に並べた擬似クリーク (上位 20) と属する企業の業種 括弧内は企業件数を表す。

順位	サイズ	業種									
1	13	繊維機械	化工機械 (1)	輪受 (3)	自動車 (2)	自動車部品 (3)	自動車関連	自動車 (2)	車体・他		
2	12	繊維二次加工	他化学	化合機 (2)	他繊維 (3)	繊維二次加工 (3)	繊維機械	総合商社			
3	12	他ゴム製品	輪受	他機械	他電気機器	自動車関連 (2)	車体・他	自動車部品 (2)	自動車 (2)		
4	11	化粧料	肥料	大手化学	塗料・ソーダ	他化学 (4)	炭素一貫	他製薬業	総合商社		
5	10	運輸関連	石灰石油製品 (4)	他金属製品 (2)	特産品	工作機械	自動車関連	自動車	建材商社		
6	10	倉庫	他化学 (3)	石油化学	塗料 (2)	他製薬業	総合商社	食品会社			
7	10	他製薬	除染	外航・他	石油精製及び販売 (2)	他商社 (2)	石炭石油製品	化工機械	総合商社		
8	9	大手海運	他商社	塗料	大手製糖	他パルプ・紙	石油精製及び販売	石炭石油製品	化工機械	総合商社	
9	9	調味料	油脂・売卸	塗料 (2)	化粧品・歯磨	農薬・殺虫剤	他金属製品 (2)	総合商社			
10	9	他化学	他建設	他サービス業	プラスチック	工作機械	総合商社	造船 (2)	総合商社		
12	8	他サービス業	機械金属商社	小売業・紙	他製薬業スタッフ	他機械	電機関連商社	自動車関連	自動車		
13	8	電機関連商社	内燃機	自動車	自動車部品 (3)	自動車部品	建材商社	自動車部品	建材商社		
14	8	自動車部品	内燃機	化工機械	自動車関連	自動車	自動車部品	自動車部品	建材商社		
15	8	自動車部品 (3)	機械金属商社	内燃機	自動車関連	自動車	自動車	自動車	自動車部品		
16	8	他輸送用機器	自動車部品 (2)	鋳造鋼	輪受	自動車 (2)	車体・他	建材商社			
17	8	内燃機	自動車関連	自動車 (3)	自動車部品	時計	建材商社				
18	8	他化学 (2)	肥料	塗料・ソーダ	塗料	他サービス業	化工機械	総合商社			
19	7	他化学	自動車	自動車 (2)	自動車部品 (3)						
20	7	自動車	車体・他	自動車部品 (4)	機械金属商社						

対象データに対して、閾値を 0.8 より 0.1 刻みで 0.5 まで変化させて抽出された擬似クリークのサイズの頻度分布を図 1 に示す。本節では、クリークのサイズをある程度大きく保ってコミュニティ理解を図る意図から閾値 0.6 を採用し、そのときの擬似クリークを分析対象とした。得られた擬似クリークの数 6,038 個において、擬似クリークのサイズの最大値は 13 が得られた。

表 3 は擬似クリークサイズの上位 20 個に含まれる企業の業種を示す。最もサイズの大きい擬似クリークに自動車業と化学業に属する企業が抽出された。またこれまで同様、商社が上位によく現れる結果となっている。ここで、擬似クリークによる企業コミュニティでは、極大クリークでは上位に現れなかった繊維業 (表 3 の第 2 位)、鉱業 (表 3 の第 8 位)、海運業 (表 3 の第 9 位) 等、多様な業種が得られている。いずれの企業も仕入・仲介・販売や物流、生産・加工・販売等企業の活動に即した企業コミュニティが得られており、企業の取引構造の理解に繋がる結果が得られていると判断できる。

4.3 結束度による取引関係の構造分析

本節では、擬似クリークに基づき得られた企業コミュニティを用い、結束度に基づき企業の取引構造の定量化について評価する。図 2 は前節で抽出した擬似クリークにより、各企業が擬似クリークのサイズ毎に含まれる頻度と各企業の結束度を重ね合わせた結果であり、各企業が擬似クリークに含まれる頻度合計の上位 20 社を左から属する業種により示す。図 2 からは、結束度の高い業種として自動車業界が、低い業種として商社や卸・小売業等が挙げられる。結束度において高

い値および低い値を示した企業のグラフ構造の可視化結果を、それぞれ図 3 および図 4 に示す。これらの図より、結束度は企業の取引が相互に依存する割合を表しており、互いに取引し合う密型となるか、取引関係が特定企業が中心に行われる星型となるかの構造に関してタイプ分類に繋がる量となっていることがわかる。

図 3 で示されているような密型は自動車業に顕著であり、部品の仕入先や販売先など業務プロセス上で密接に協調して生産を行っていることが取引構造に反映されていることが示唆される。こうした取引構造では、一旦どこかの企業がサービス提供できない状況になると悪影響が全体へ波及するというリスクを持つため、高い結束度を持つ企業に対する与信では、取引先の取引先のようにコミュニティを捉えた評価が有効と考えられる。

図 4 で示されているような星型は商社や卸売業など仲介を行う企業に顕著であり、中心には仲介する企業が、取り巻きには仕入先・取引先となる企業が位置づけられる取引構造となっている。こうした取引構造では、円滑に商品の仕入・販売ができるが、売上不振になると取引が停止されるというリスクも持つことが知られる。また取引に当たっては、与信評価を行う必要性があるが、バックボーンとなる企業の存在がないため、与信評価は従来通りの取引先のものに着眼することで構わないと判断し得る。

以上のように、結束度として定義した指標の大小に応じて、企業の取引構造を密型や星型にタイプ分類することができ、企業のコミュニティにおける活動形態

☆ 物流系企業では、売掛金の回収等リスク審査が経営上の課題となっている⁷⁾

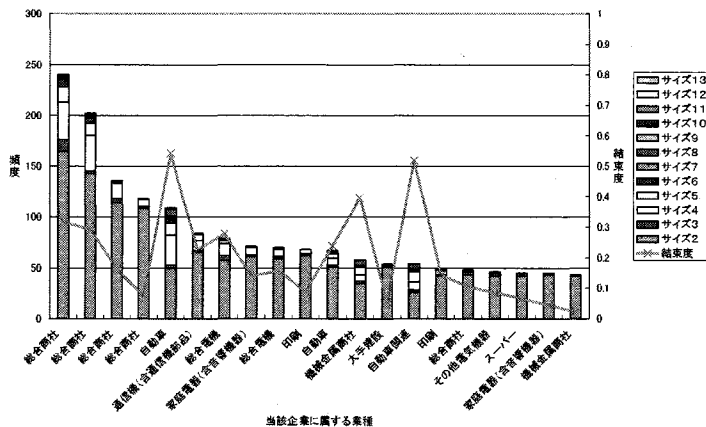


図 2 疑似クリークに含まれる頻度上位 20 社におけるサイズの構成と結束度

の理解と共に企業評価への活用可能性が広がる。

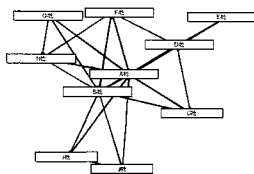


図 3 結束度の値が大きい企業の取引構造 (密型) の例

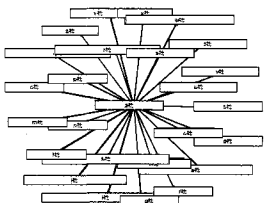


図 4 結束度の値が小さい企業の取引構造 (星型) の例

5. ま と め

本論文では、極大クリーク全列挙アルゴリズムを用いた企業コミュニティ抽出について述べ、クリークを用いて関係構造の特徴をタイプ分類する結束度を考案した。考案した指標を用いた実データに対する実験

により、互いに取引し合う密型と特定企業を中心に取引が行われる星型とのタイプの定量化を可能にし、クリーク抽出が企業コミュニティ理解に有用であることを示した。

謝辞 本論文における実験に用いた企業データを提供して頂いた日本経済新聞社に感謝いたします。

参 考 文 献

- [1] Barabási, A.-L.: *LINKED: The New Science of Networks*, Perseus Book Group 2002.
- [2] 安田 雪, 自動車部品サプライヤーの取引構造にみる「べき法則」, *University of Tokyo MMRC Discussion Paper*, No. 33, 2005.
- [3] 米森力, 松永務: グラフスペクトル法を用いた企業コミュニティ抽出, *日本ソフトウェア科学会, JWEIN '06*, 2006.
- [4] E. Tomita, A. Tanaka and H. Takahashi: The worst-case time complexity for generating all maximal cliques and computational experiments, *Teoret. Comput. Sci.* 363, pp. 28-42, 2006.
- [5] K. Makino and T. Uno, New algorithms for enumerating all maximal cliques, *Proc. of 9th Scandinavian Workshop on Algorithm Theory*, Denmark, 2004.
- [6] M. Haraguchi and Y. Okubo: A method for pinpoint clustering of web pages with pseudo-clique search, *Federation over the Web, International Workshop, Dagstuhl Castle, Germany, May 1 - 6, 2005, Revised Selected Papers*, Springer-LNAI 3847, pp. 59-78, 2006
- [7] G. エメリー, 有賀健, 河口品彦: 企業間信用と企業間関係の分析, 大蔵省財政金融研究所 フィナンシャル・レビュー, 1993.