

取引ネットワークに依拠した企業の振る舞いの分析 －ソフトウェア産業における現状

井 上 寛 康

Analysis of Firms' Behavior Based on Transaction Networks: Current State of Software Industry

INOUE Hiroyasu

Abstract

Firms have to provide for international competition more than ever before. To do this, they have to exploit resources around them so that they can set off deficiency of ability. Therefore, they try to build relations with other firms. There are various types of relations, such as transactions, shareholdings, executive dispatching, cooperative researches, developments, and alliances. In addition, the relations express complex networks. This paper treated transaction networks as one of such networks and selected the Japanese software industry as an example. This was because it has been pointed out that the industry has some structural defect and a whole picture of the structure has not been revealed. I used a p^* model to analyze the structure.

キーワード：企業，ネットワーク，取引，ソフトウェア

Keywords : Firm, Network, Transaction, Software

目 次

1. はじめに
2. p^* モデル
3. 分析
4. 結論

1 はじめに

企業は競争の国際化によって、従来以上に生き残りのための競争力を要求されている。より高い競争力を実現するにあたっては、企業が提供できる財には限りがあるため、企業内部を注視するのみならず、その企業を取り巻く周囲の状況に対応する必要性が非常に高まっている。そこで企業は、自社の能力を最大限に活かすために、周囲との関係性を適切に構築しているとされている。具体的に企業間関係性は、取引、株所有、役員派遣、共同研究、アライアンスなど、多く存在する。また、それらは複雑に絡み合ったネットワークを形成している。

本論文では、そのような現在の産業構造において、特に取引ネットワークを取り扱う。そして、パイロット的なサンプルとして、日本のソフトウェア産業を取り上げる。

日本のソフトウェア産業は多くの構造上の問題が指摘されている。その1つに建設業で見られるような、受注内容を丸投げする階層的な構造がある。このような構造は、不必要な価格の上昇や下請けの搾取が起るなど弊害が多い。このような問題点が指摘されながらも、このネットワーク上の構造が明確に把握されたことはいまままでない。

本論文では p^* モデル^[1]による分析を行う。この分析手法はネットワークに現れる構造のうち、特に意味のあるものを洗い出す手法である。

2 p^* モデル

p^* モデルとは、我々の社会における関係性の網（ネットワーク）において、ある人（あるいは物）とある人（物）の関係が、他の（どのような人の関係であるかを問わない）関係にどのように影響を与えるのか、その背後にあるモデルをあぶりだすものである。端的にいえば、どのような局所構造が現れやすいネットワークであるかを求めることである。これは関係性の関係性を求めることといえる。

p^* モデルは Markov random graph として、Frank らによって初めに議論された。 p^* モデルは今日に至るまで、多くの社会ネットワーク分析で活用されている有用な分析手法である。モデルに対する計算方法は Strauss ら^[2]によってより深い議論がなされたが、Wasserman ら^[3]によって一般的な形として完成された。多くの派生 p^* モデルがある^[3, 4]が、ここで議論するのは、Wasserman らによる解説^[5]を元にする。

2.1 表記法

表記法は Wasserman ら [6] に準ずる。社会ネットワークとは n 個のノード (アクター) を持ち、そしてそれらの間に r 個のリンク (社会的関係) が形成されているとする。ここではあるリンクが存在する場合を 1 とし、他の場合を 0 とする。多値化や多重の関係についてはさらに議論が深められている [7, 8, 9, 10, 11]。

$N = \{1, 2, \dots, g\}$ をノードの集合とする。 X をノードのペアを要素とする集合とする。したがって、リンクは $g \times g$ の行列 X として次のように表すことができる。

$$X_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \mathcal{X} \\ 0 & \text{otherwise} \end{cases}$$

X とその要素はランダムであるとする。すなわち、あるリンクが存在するかどうかは確率的に決まると仮定している。この確率的であるという仮定を元にして、リンクとリンクの間には関係、すなわち p^* モデルを求めることにつながる。この p^* モデルは依存グラフにより定義されるが、これについては後の節で述べる。

p^* モデルの説明は段階を経て行すが、その準備として、 \mathbf{X}_{ij}^+ 、 \mathbf{X}_{ij}^- および \mathbf{X}_{ij}^c を定義する。

$$\begin{aligned} (\mathbf{X}_{ij}^+)_{mn} &= X_{ij}, \text{ if } (m, n) \neq (i, j) \\ &= 1, \quad \text{if } (m, n) = (i, j) \end{aligned}$$

$$\begin{aligned} (\mathbf{X}_{ij}^-)_{mn} &= X_{ij}, \text{ if } (m, n) \neq (i, j) \\ &= 0, \quad \text{if } (m, n) = (i, j) \end{aligned}$$

$$\begin{aligned} (\mathbf{X}_{ij}^c)_{mn} &= X_{ij}, \text{ if } (m, n) \neq (i, j) \\ &= \text{undefined}, \quad \text{if } (m, n) = (i, j) \end{aligned}$$

2.2 依存グラフによる説明

前節で p^* モデルでは各リンクは確率変数であるとした。あらゆる統計モデルに必要なことは、その確率変数の間の統計的な依存関係を知ることである。そのために必要な道具として依存グラフを説明する。この依存グラフは Hammersley-Clifford の定理 [12] を適用する前提となる。

あるネットワーク $x = [x_{ij}]$ は確率変数の集合 $X = [X_{ij}]$ の1つの例に過ぎないとする。これら確率変数集合の要素がすべて独立であると考えるのは極めて不自然であるから、それらの間の関係を考えることになる。それが依存グラフである。

依存構造を表した依存グラフ D を確率変数集合 X の間の依存関係であるとする。依存グラフ D はノードとして、確率変数集合 X における（元のネットワークのノードの）インデックス $\{(i, j); i, j \in N, i \neq j\}$ に対応する要素を持つ。そして、それらの任意の2つのノードの間のリンクは条件付従属とする。このような考え方はグラフィカルモデリングと共通である。グラフィカルモデリングの社会ネットワークへの適用については研究が多くある^[13, 14, 15]。

ネットワークの分布には大きく分けて3つのクラスがあるとされている^[16, 17]。それらは、ベルヌーイグラフ^[10]と条件単一グラフの分布、2項間依存分布^[18, 19]と p^* である。ほかにも存在する^[20]がこれは数学的な研究であり、ノードが無限に存在するときの漸近的な振る舞いについて調べている。前述の3つのクラスについて順に説明する。

ベルヌーイグラフでは、異なるリンクはすべて独立であるとする。すなわち、 X_{ij} と X_{kl} は $i \neq j$ かつ $k \neq l$ である限り独立である。このような分布に対しては、依存グラフはリンクを持たない。したがってあるリンクの確率が単に P_{ij} によって決まる。これがベルヌーイと呼ばれる理由である。特にこのグラフにおいて確率を同じとしたのを条件単一グラフと呼ぶ。このグラフにおいては、すべてのリンクは独立でかつ同じ生起確率を持つ。条件単一グラフにはいくつかのバリエーションがある。たとえば、依存グラフにおいて指定したノードのみがある生起確率 l を持ち、それ以外は0であるグラフなどがある。

2項間依存分布は、 X_{ij} と X_{kl} は $(k, l) = (j, i)$ のみ依存関係にあるとする。言い換えれば任意の2組の2項間の間は独立である。このとき、依存グラフは2つのノードの間のみリンクを持つ。すなわち、 $\varepsilon_D = \{(i, j), (i, k), \text{ for all } j \neq k \text{ for every } i\}$ のみである。このモデルは p_1 モデルという言い方もされている^[19]。リンクの生成確率を変数として捉え、それらの間に依存関係があることを考えるのが自然であるが、ここまでの2つのモデルはそれを表現できない。あまりに限定的なモデルといえる。

ここまでの議論を踏まえ、一般的な分布とできるのが p^* である。これは依存グラフにおける任意のノードの間にリンクを仮定する。 p^* の基本的なモデルでは、依存グラフの X_{ij} と X_{kl} において、 $\{i, j\} \cap \{k, l\} = \emptyset$ ならばリンクが存在しない。すなわちある1つのノードが関係したリンクの間でのみ、依存関係があるとするモデルである。これはマルコフ空間処理と同じであり、最初はこのモデルをマルコフグラフ^[1]としていた。

ここまでクラスの議論をしてきた一方で、当然であるが任意のリンクを持つ依存グラフ

を考えれば最も一般的なクラスとなる。しかし、そのような依存グラフは求め得ない。これは一般的に実際のデータが1つしか得られないことと、求めようとする関係性の数 ($O(n^2)$) が多すぎることによる。

2.3 p* モデルの定義

あるネットワーク x は確率変数の集合 $X = [X_{ij}]$ の1つの例に過ぎないとし、その X には背後に依存グラフ D が存在すると仮定することについてこれまで説明してきた。p* モデルではこの依存グラフにおけるクリーク (完全部分グラフ) を取り扱う。クリークとは、依存グラフのノード N_D の部分集合が、すべてリンクで結ばれている部分グラフのことをさす。したがって、そのクリークにおいては元のネットワークのリンクにおいてすべて従属関係にあることを指す。

Hammersley-Clifford の定理^[12] は X の確率モデルは依存グラフ D のクリークにのみ依存していることを示した。それによりあるネットワーク $Pr(X = x)$ を次のように表した。

$$Pr(\mathbf{X} = \mathbf{x}) = \kappa^{-1} \exp\left(\sum_{A \subseteq N_D} \lambda_A \prod_{(i,j) \in A} x_{ij}\right) \quad (1)$$

ここで

- $\kappa = \sum_x \exp(\sum_{A \subseteq N_D} \lambda_A \prod_{(i,j) \in A} x_{ij})$ は正規化定数。
- D は依存グラフであり、式のサンメンションは全ての部分集合であることを示している。
- $\prod_{(i,j) \in A} x_{ij}$ は十分統計量である。十分統計量とは「ある統計データに対し、それが従う確率分布を示す母数 θ に対応する統計量の値が決められた条件下で、データが出現する条件付き確率分布が、もはや θ にはよらない場合に、この統計量は十分である」^[21] という意味である。
- D の部分グラフ A がクリークでないならば $\lambda_A = 0$ である。

結局のところ、 $\lambda_A \neq 0$ なのは x の最大クリーク (クリークの中でノード数最大のもの) さえ考えればよい。最大クリークは他の全てのクリークを含むからである。

前述したが、この λ_A は明らかに多すぎる。したがって一様性、あるいは同形性といってもよいが、同じ形のクリークについては1つのパラメータで表してしまうことにする。たとえば、2つのノードが関係する (ダイアド) 場合は、それらの間で無関係、一方向、双方向のリンクがありうる。一方向と双方向は、それぞれ依存グラフのノードが1つと2

つの同形完全部分グラフと考えられる。また、3つのノードが関係する（トライアド）の場合は、15種類の同形完全部分グラフがある。

この同形性を考慮してもまだパラメータは多すぎる。なぜならいくらかでも複雑な依存グラフが考えられるからである。そこでたとえばトライアドよりも多い（4つ以上のノードが関係する）場合は $\lambda_A = 0$ とする必要がある。こういった高次のものは解釈も難しくなる。

2.4 計算の方法

式（1）は次のような形に書くことができる。

$$Pr(\mathbf{X} = \mathbf{x}) = \frac{\exp\{\theta' \mathbf{z}(\mathbf{x})\}}{\kappa(\theta)} \quad (2)$$

ここで

- θ はパラメータのベクトルである。
- $\mathbf{z}(\mathbf{x})$ はネットワークの統計量である。
- $\kappa(\theta)$ は正規化定数である。

この式はロジスティック回帰式と見ることが出来る^[3]。この θ を求めるという点から見ても妥当といえる。

尤度関数は単純であり、次のように表される。

$$L(\theta) = \frac{\exp\{\theta' \mathbf{z}(\mathbf{x})\}}{\kappa(\theta)} \quad (3)$$

このように単純ではあるが、求めることは容易ではない。これは $\kappa(\theta)$ が未知あることによる。厳密な対数尤度を求めることはできなくはないにしても極めて困難である。中にはそのような研究もある^[22] が、ごく小さなネットワークに限られる。

最近用いられている方法には2つある。（1）最大擬似尤度算出法^[23, 24, 25, 26, 2]と（2）マルコフ連鎖モンテカルロ最大尤度算出法^[27, 28]である。

2.5 最大擬似尤度算出法

確率の計算には式（2）における κ が計算されなければならないが、極めて難しい。そこでそれを回避する方法が必要となる。ここではオッズ比を用いるロジスティック回帰に

よる方法について述べる。

式 (2) をロジスティック回帰式とする。これにより概算の尤度を求めることができる。まず, X_{ij} の補集合を用いて, X_{ij} について分割する。

$$\begin{aligned} P(X_{ij} = 1 | \mathbf{X}_{ij}^C) &= \frac{Pr(\mathbf{X} = \mathbf{x}_{ij}^+)}{Pr(\mathbf{X} = \mathbf{x}_{ij}^+) + Pr(\mathbf{X} = \mathbf{x}_{ij}^-)} \\ &= \frac{\exp\{\theta' \mathbf{z}(\mathbf{x}_{ij}^+)\}}{\exp\{\theta' \mathbf{z}(\mathbf{x}_{ij}^+)\} + \exp\{\theta' \mathbf{z}(\mathbf{x}_{ij}^-)\}} \end{aligned}$$

これによってオッズ比を表すと,

$$\begin{aligned} \omega_{ij} &= \log \frac{Pr(\mathbf{X}_{ij} = 1 | \mathbf{X}_{ij}^C)}{Pr(\mathbf{X}_{ij} = 0 | \mathbf{X}_{ij}^C)} \\ &= \theta' [\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)] \end{aligned} \quad (4)$$

となる。

$\delta(x_{ij}) = [\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)]$ とするならば, 式 (4) は $\omega_{ij} = \theta' \delta(x_{ij})$ と表せる。この $\delta(x_{ij})$ は x_{ij} の 0 と 1 でどれぐらいネットワーク統計量が変わるかを表す。条件付独立を仮定することによってオッズ比は尤度の概算に用いることができる。

この概算には次の関数を用いる。

$$PL(\theta) = \prod_{i \neq j} Pr(X_{ij} = 1 | \mathbf{X}_{ij}^C)^{x_{ij}} Pr(X_{ij} = 0 | \mathbf{X}_{ij}^C)^{1-x_{ij}} \quad (5)$$

最大擬似尤度指標は式 (5) を最大化するような θ である。このアプローチはリンクの間の条件付独立性を仮定している。 θ の算出はロジスティック回帰によって求まる。式 (5) の擬似尤度を最大化することと, 式 (4) にロジスティック回帰を行うことと同じである。

ロジスティック回帰分析を用いられることから, この方法は極めて簡単といえる。すなわち, すべての X_{ij} をケースと考え, 依存グラフにより表される独立変数を求めることになる。

2.6 最大尤度算出法

前述の擬似尤度算出法は Besag により空間統計学の分野で開発された^[23]。この方法にはいくつも利点があり, 複雑なモデルでも比較的簡単に求められる。それにもかかわらず擬似尤度算出法はあまり理解されなかったため, 最大尤度をモンテカルロ法によって推測

する方法が発達した。

モンテカルロ法を用いた算出にはいくつものバリエーションがある。基本的な考え方はパラメータセットを適当に用意して、シミュレーションした結果それを手直しするというものである。そしてこれをある程度の安定が得られるまで繰り返すことになる。これには2つのメジャーな方法があり、1つはギブスサンプラー^[29]であり、もう1つはメトロポリスハスティングスアルゴリズム^[30]である。実は後者は前者を含むことがわかっており、これを元にして、モンテカルロ法による最大尤度算出法の一般的なものはGeyerらによって提案された^[31]。Crouchらはこの方法がいかにか p^* モデルに適しているかを示した^[27]。

モンテカルロ法において重要な問題点が明らかになった。それは縮退であり、一種の局所解である。依存グラフが全ノードについてリンクを持つか、まったくリンクがないかのいずれかに陥ってしまうことがある。Strauss^[26]は最初にメトロポリスハスティングスアルゴリズムをマルコフランダムグラフに用いたが、そのときに縮退を発見した。その後につづく研究^[32, 33, 34]はいかなる方法によってもこの p^* モデルのモンテカルロ法による縮退の問題は解き得ないと述べた。

2.7 最大擬似尤度算出法と最大尤度算出法の比較

それでは結局どちらがよいかということであるが、最大尤度算出法ならば、統計的に正しい結果であることは間違いなく、予想に際しての誤差についても正しく求まる。これは最大擬似尤度算出法にはできないことである。一方で、最大尤度算出法は極めて計算コストが高く、大きなネットワークでは現実的な時間で求めることができない。そこで、最大擬似尤度算出法はいつなら適用可能か考える必要がある。

実際にこれを比較した先行研究がある^[35]。結論は非常にシンプルである。40ノードより小さければ最大尤度算出法が最大擬似尤度算出法より優れている。40から100ノードではあまり変わらないとしている。

最大擬似尤度算出法は結局のところ概算にすぎない。したがって可能なら最大尤度算出法を用いたほうがよい。最大擬似尤度算出法の使いどころがあるとすれば、 p^* モデルによって求められるパラメータが正確性ではなくておおよその傾向をつかむ場合にのみ使われる場合である。一般的にはこれで十分なはずである。逆に厳密さが求められる場合には最大擬似尤度算出法を用いるべきではない。あらゆる統計量、たとえば誤差の検定などは、最大擬似尤度算出法ではまったく効果がない。逆に最大尤度算出法は厳密ではあるが、前述の縮退の問題がある。3つのノードの構造を明らかにするのはおそらく問題ないが、それ以上大きい場合は縮退する可能性が強くなる。

3 分析

3.1 データ

本論文では、東京商工リサーチ社のデータ（TSR データ）を用いる。このデータは、約80万社の財務情報、事業情報に加えて、企業間の取引情報を含んでいる。ソフトウェア産業に該当するのは、日本標準産業分類における39情報サービス業内の、391ソフトウェア業内の、3911受託開発ソフトウェア業および3912パッケージソフトウェア業である。本業によって企業を分けた場合、3911と3912のいずれかに属する企業で、かつそれらの中で取引を行っている企業は、TSR データに4,938社あった。しかしながら、この行列データは計算機で扱うには少し大きすぎたため、3911に絞った。この場合は3,835社ある。

受託開発ソフトウェア業内の取引ネットワークに限ったということは、もし請負った企業がその自社のみで開発を完了している場合、ネットワークには1つのリンクも存在しないことになる。逆に、リンクが存在するということは、受託開発ソフトウェア業同士の取引であるから、下請けの構造が存在することになる。

図1は、企業間のリンクの例を示している。左がサービスや製品の提供元であり、右がその提供先である。金の流れとは逆であるので注意する必要がある。

図2は、受託ソフトウェア企業取引ネットワークである。矢印の方向とともに示されている。この図は UCINET^[36] に含まれている Pajek を用いて描画した。ここで用いられている描画方法は、Kamada-Kawai の方法^[37] であり、リンクをばねのように見立てて位置を決定する。ノードが持つリンクの数を次数と呼ぶが、この方法だと描画の中心部は次数の大きいものが集まる傾向にある。中心部の様子がわかりにくいのはそのためである。このような大規模ネットワークの描画は常に一長一短がある。仮にノードを円形に配置すると、リンクで円内が真っ黒な図になり、この場合はネットワークの様子が一切わからなくなる。



図1：企業間リンク例

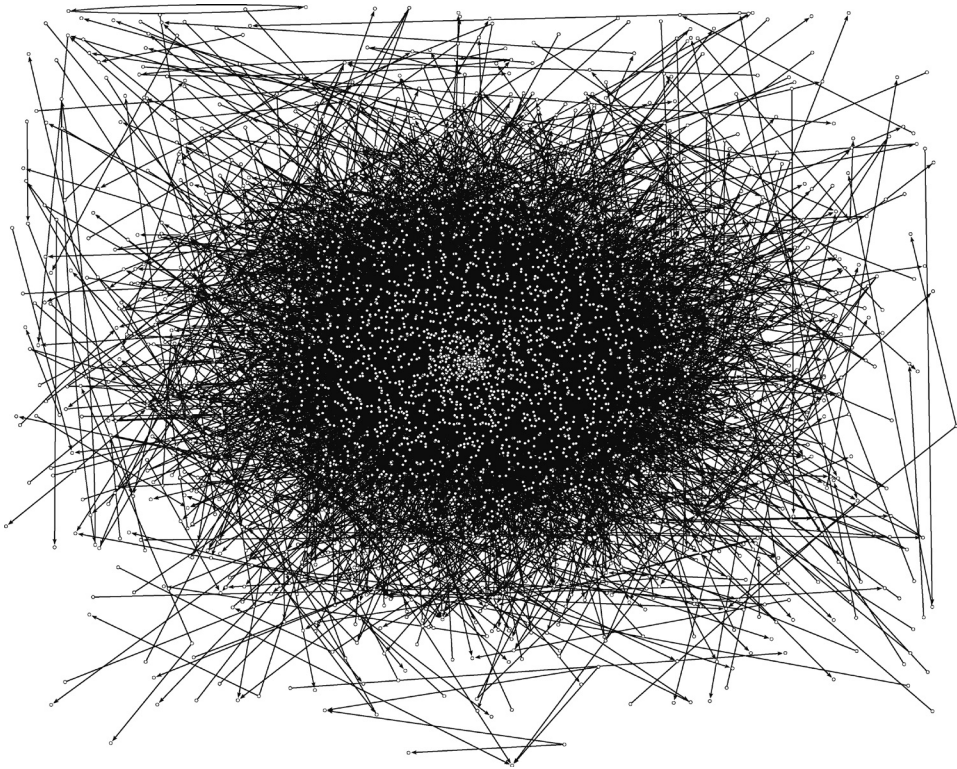


図2：受託ソフトウェア企業取引ネットワーク

3.2 分析した構造

図3は、今回分析した構造である。p*モデルによる分析は、このような部分的な構造が現れやすいかを調べることであると考えるとわかりやすい。

上から Choice, Reciprocity, Transitivity, Cycle, Two star in, Two star out, Two star mix である。特に Choice は、あるリンクが発生する確率はそのリンクの存在自身に影響する、ということで、重回帰分析上は定数項のみ、あるいは説明変数なし、という状況のことである。

取引ネットワークに依拠した企業の振る舞いの分析 (井上寛康)

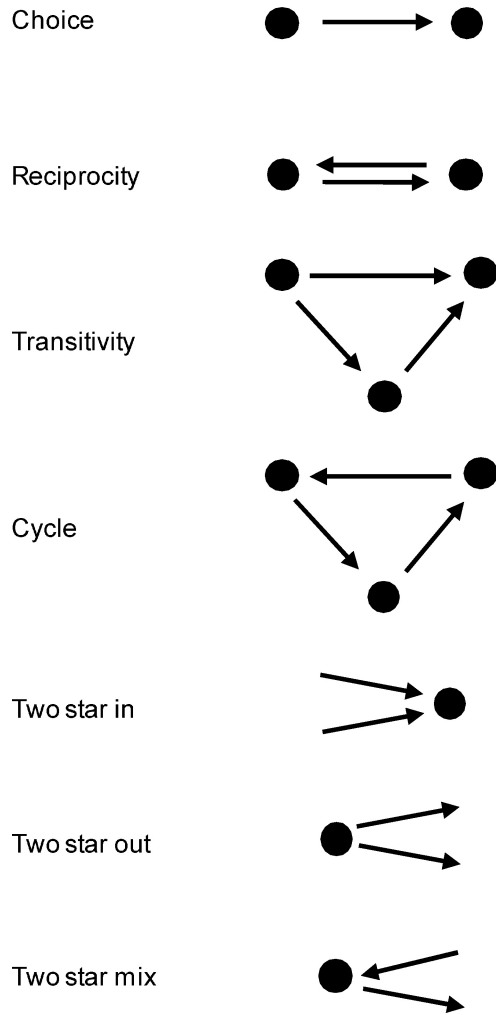


図3：分析した構造

表1：モデル回帰結果

Model	G_{PL}^2	Variables	AIC	Difference to 1
1. Choice	113,165.0	1	113,163.0	-
2. Reciprocity	109,308.1	2	109,312.1	3,856.9
3. Transitivity	113,011.5	2	113,015.5	153.5
4. Cycle	113,051.6	2	113,055.6	113.4
5. Two star in	107,191.8	2	107,195.8	5,973.2
6. Two star out	112,692.6	2	112,696.6	472.4
7. Two star mix	112,663.1	2	112,667.1	501.9

3.3 擬似尤度法による分析結果

前節で述べた擬似尤度法をロジスティック回帰分析により行った。この分析には SPSS 17.0を用いた。表1 は分析結果を表している。

受託ソフトウェア産業ネットワークに対して、各モデルに対する -2 対数尤度 (G_{PL}^2) およびそのときの説明変数の数、赤池情報量基準 (AIC)、 -2 対数尤度を比較したモデル1と比較した差を示している。

3.4 議論

表1のモデル1. Choice は、前述したように定数項のみで回帰させた場合である。その際の -2 対数尤度ほかが表示されている。これが他の分析の基準となる。

結果を見ると、モデル2. Reciprocity とモデル5. Two star in は突出してその効果が大きいことがわかる。モデル2. Reciprocity は、ある2つの企業の間で、片方向に取引があると、反対方向にも取引を生じやすいか、を示している。モデル1との -2 対数尤度の差は3,856.9である。ソフトウェア産業では下請け関係が強調される傾向にあるが、実際にはこのような双方向性が存在することは特筆すべきである。またモデル5. Two star in は他の1社からサービス/製品を購入する、すなわち発注を行うと、それは2社以上を行う傾向にあることがわかる。

モデル3. Transitivity, モデル4. Cycle は、分析されたモデルの中で3つのノードが関係するものであるが、その影響はあまりないことがわかる。つまり取引があるとしても、ソフトウェア産業では3社が関係するのは非常に少なく、2社の関係が基本であることがわかる。

モデル6. Two star out, モデル7. Two star mix は、5. Two star in と同様に2つの関係を扱ったモデルであるが、これらの影響は非常に少ないとわかる。モデル6. Two star out が少ないということは、下請けを行う会社は2社以上ではなく、1社と専属的に取引を行う傾向が強いことを示す。またモデル7. Two star mix が少ないということは、下請けをしてさらに発注するような傾向は少ないことがわかる。日本のソフトウェア産業は、建設業に見られるような階層的な下請けの構造を持つとこれまで考えられていたが、以上の分析は、これとは一致していない。もし階層的な下請けが多く存在すれば、モデル7. Two star mix が現れやすいはずである。すなわち、階層性はないか、あまり強くないことがわかる。

次に、モデル2. Reciprocity が強く現れていることから、単純に下流に流れていくような取引の構造ではないということである。ただし、Reciprocity は取引内容や金額によら

ないので、本業のソフトウェアでない、あるいは小額の他の取引なども同等に扱われるので、今後より深い調査が必要となる。

受託ソフトウェア企業取引ネットワークを用いた本論文の分析で初めて明らかとなったことの1つに、 p^* モデルが得意とする分析のあり方である。今回のネットワークはノード、すなわち企業の働きがいくつかに分かれていた。これは、発注のみ、受注のみ、のようにリンクの特徴が分かれていたということである。このような役割の違いを社会ネットワーク分析ではモードと呼ぶが、暗黙的なモードが存在する場合、全体を一様に評価する p^* モデルでは分析が難しくなる側面がある。人の友人ネットワークに代表されるようなフラットな構造の場合、分析の効果がやすいといえる。

また、これまでノードの数はせいぜい100程度とされてきた p^* モデルによる分析を、数千のノードで行ったことから、そのフィージビリティに対する検証も行ったことになる。 p^* モデルは基本的に行列の演算があるため、計算量が時間および空間において $O(n^2)$ になる。この計算が可能な計算機は、現在ではかなり高い能力となる。

一般的に回帰モデルは説明変数が増えるほど G_{PL}^2 が小さくなるので、その有意な差を決める必要がある。これは $-2n(n-1)\log(1-\delta)$ のように定義される。ここで n はノードの数、 δ は定数であり、一般的に0.001や0.005が用いられる。ここでは0.001を用いるが、この場合は有意差は12777.6になる。今回のデータでは、ほとんどの行列要素が0となるために、ロジスティック回帰で得られる係数だけでなく、尤度も小さくなりがちである。したがってこのような大きな尤度の差はここでは得られなかった。その理由は、モデル1の尤度が113,165.0であり、現実的には起こりえない数値となるためである。このような尤度の差による判定は、大規模なネットワークには適切ではないといえる。

本論文の問題点は交互作用を考慮していないことである。理由は2つあり、1つは交互作用を発見した場合に、今回のようなケースにおいてどのように解釈するのか、はっきりしないためである。もう1つは手法上の制約であり、これは擬似尤度を算出しているに過ぎないため、交互作用の発見自体にそもそもどの程度意味があるのかははっきりしないためである。

4 結論

本論文では、産業構造の1つの把握方法として、企業ネットワークを分析した。そのようなネットワークとして、日本のソフトウェア産業における取引ネットワークを取り扱った。これは、日本のソフトウェア産業は多くの構造上の欠点が指摘されているが、その全

貌が把握されたことはまだないためである。

p*モデルによる分析により、建設業的とされたソフトウェア産業の常識とは異なる、以下のような結論を得た。1. サービスや製品の相互のやり取りがある。これは、単に上流から下流へと流れる下請けだけの木構造ではなく、より複雑なネットワークの構造を持っている。2. 下請けを行う会社は1社の専属的な傾向が強い。3. ソフトウェアの下請けも発注も行う会社は少ない。4. ソフトウェアの発注を行う会社は複数の会社から行う傾向が強い。

謝 辞

本研究は科研費（20730268）の助成を受けたものである。

参考文献

- [1] O. Frank and D. Strauss. Markov graphs. *Journal of the American Statistical Association*, 81 (395) :832-842, 1986.
- [2] D. Strauss and M. Ikeda. Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*, 85:204-212, 1990.
- [3] S. Wasserman and P. Pattison. Logit models and logistic regressions for social networks: I. an introduction to markov random graphs and p*. *Psychometrika*, 60:401-426, 1996.
- [4] C. J. Anderson, S. Wasserman, and B. Crouch. A p* primer: Logit models for social networks. *Social Networks*, 21:37-66, 1999.
- [5] S. Wasserman and G. Robins. An introduction to random graphs, dependence graphs, and p*. In P. J. Carrington, J. Scott, and S. Wasserman, editors, *Models and Methods in Social Network Analysis*, pages 148-161. Cambridge University Press, 2005.
- [6] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
- [7] P. Pattison and S. Wasserman. Logit models and logistic regressions for social networks: Ii. multivariate relations. *British Journal of Mathematical and Statistical Psychology*, 52:169-193, 1999.
- [8] G. L. Robins, P. Pattison, and S. Wasserman. Logit models and logistic regressions for social networks, iii. valued relations. *Psychometrika*, 64:371-394, 1999.
- [9] O. Frank. Statistical analysis of change in networks. *Statistica Neerlandica*, 45:283-293, 1991.
- [10] O. Frank and K. Nowicki. Exploratory statistical analysis of networks. In J. Gimbel,

- J.W. Kennedy, and L.V. Quintas, editors, *Quo Vadis Graph Theory? A Source Book for Challenges and Directions*. Amsterdam, 1993.
- [11] O. Frankl. Composition and structure of social networks. *Mathematiques, Informatique, et Science Humaines*, 137:11–23, 1997.
- [12] J. E. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36 (192-236) , 1974.
- [13] S. Lauritzen. *Graphical Models*. Oxford University Press, 1996.
- [14] G. L. Robins. Graphical modelling. *Chance*, 10:37–40, 1997.
- [15] G. L. Robins. *Personal Attributes in Inter-personal Contexts: Statistical Models for Individual Characteristics and Social Relationships*. PhD thesis, University of Melbourne, 1998.
- [16] P. Pattison and S. Wasserman. Social network models, statistical. In *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier Science, London, 2001.
- [17] S. Wasserman and P. Pattison. Statistical models for social networks. In H. Kiers, J. Rasson, P. Groenen, and M. Schader, editors, *Studies in Classification, Data Analysis, and Knowledge Organization*. Springer-Verlag, Heidelberg, 2000.
- [18] S. Wasserman. Conformity of two sociometric relations. *Psychometrika*, 52:3–18, 1987.
- [19] P. W. Holland and S. Leinhardt. An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, 76:33–65, 1981.
- [20] B. Bollobas. *Random Graphs*. Academic Press, London, 1985.
- [21] 十分統計量の項 . wikipedia. <http://ja.wikipedia.org/wiki/>.
- [22] M. E. Walker. *Statistical Models for Social Support Networks: Application of Exponential Models to Undirected Graphs with Dyadic Dependencies*. PhD thesis, University of Illinois, Chicago, 1995.
- [23] J. E. Besag. Statistical analysis of non-lattice data. *The Statistician*, 24:179–195, 1975.
- [24] J. E. Besag. Some methods of statistical analysis for spatial data. *Bulletin of the International Statistical Association*, 47:77–92, 1977.
- [25] J. E. Besag. Efficiency of pseudo-likelihood estimation for simple gaussian random fields. *Biometrika*, 64:616–618, 1977.
- [26] D. Strauss. On a general class of models for interaction. *SIAM Review*, 28:513–527, 1986.
- [27] T. A. B. Snijders. Markov chain monte carlo estimation of exponential random graph models. *Journal of Social Structure*, page 2 , 2002.
- [28] M. S. Handcock. Statistical models for social networks: Inference and degeneracy. In R. Breier, K. Carley, and P. Pattison, editors, *Dynamic Social Network Modeling and Analysis*, pages 229–240. National Academies Press, Wahington, DC, 2003.
- [29] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,

6:721-741, 1984.

- [30] S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *American Statistician*, 49:327-335, 1995.
- [31] C.J. Geyer and E. A. Thompson. Constrained monte carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society Series B*, 54:657-699, 1992.
- [32] J. E. Besag. Markov chain monte carlo for statistical inference. Technical report, University of Washington, Center for Statistics and the Social Sciences, Seattle, 2000.
- [33] M. S. Handcock. Progress in statistical modeling of drug user and sexual networks. Technical report, University of Washington, Center for Statistics and the Social Sciences, Seattle, 2000.
- [34] P. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97:1090-1098, 2002.
- [35] J. Corander and K. Dahmström. Maximum likelihood estimation for markov graphs. Technical Report 1998:8, University of Stockholm, Department of Statistics, Stockholm, Sweden, 1998.
- [36] S.P. Borgatti, M.G. Everett, and L.C. Freeman. *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologie, Harvard, 2002.
- [37] T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31:7-15, 1989.