

アフィリエーションネットワーク概念を用いた 組織内情報分析法のモデル化と 情報ネットワーク制御への応用

Social Network Analysis Model based on Affiliation Network and Usage toward Information Network Control

栗野俊一†
Syun-ichi Kurino

吉開範章‡
Noriaki Yoshikai

1. まえがき

企業が災害、あるいは事故に遭ったとしても、重要事業を中断せず、仮に中断した場合でも、極力短期間に再開し、顧客の他社への流出、マーケットシェアの減少、企業評価の低下等の脅威から企業を守るために、事業継続計画 (Business Continuity Plan) を作成することが経営戦略上、重要視されてきている⁽¹⁾。BCP 作成のためには、業務の現状分析の後、ビジネス機能とIT機能に分離して個々に復旧計画を立てる方法が取られる場合が一般的である⁽²⁾が、本来、企業内のビジネス機能は、企業内に構築・使用される情報ネットワーク上に展開される社員間の社会的ネットワークと融合した社会活動として実現されるものと考えられ、復旧計画を立案する場合も、ビジネスの復旧特性と、それを実現可能な情報ネットワークでの復旧特性を総合的に規定することが必要である。また重要プロジェクトあるいはキーパーソンの決定法は、復旧計画のキー要素となるが、その方法は、公にはされていない。

一方、企業内の電子メール分析を元にユーザー行動を定量的にモデル化し、組織内の業務プロセス分析を行う研究が進められている⁽³⁾が、それらは、個人のみを対象を限定し、個人が所属する組織との相互関係まで考慮した検討はなされていない。

本稿では、企業の活動状況を可視化する方法に、「組織と個人の2重性」を考慮したアフィリエーションネットワーク⁽⁴⁾を用いたモデルが有効であることを示すと共に、ケーススタディーとして、現実の組織内のメールヘッダ情報を使った業務分析を行い、BCP 作成にも有効な情報ネットワークの品質制御に応用する方法について言及する。

2. 準備

モデル化に、グラフ理論を用いるので、使用した記号と意味について説明する。

2.1 グラフ

[グラフ: G] ノード(点)集合 N と、エッジ(辺)集合 E を用いて、 $G = \langle N, E \rangle$ で表す。ノードはノードにつけられた番号 (1, 2, ..., M) で、エッジはその始点と終点の番号の順序対 $\langle i, j \rangle$ で、それぞれ表現するが、本論ではすべて、無向グラフを扱うので、この順序には意味が無い。

[隣接行列: X] グラフ $G = \langle N, E \rangle$ の隣接行列 X は、次のように定義される。

$$X = (x_{i,j}), x_{i,j} = \begin{cases} 1 & (\langle i, j \rangle \in E) \\ 0 & (\langle i, j \rangle \notin E) \end{cases}$$

[誘導される部分グラフ: G_V] グラフ $G = \langle N, E \rangle$ の部分グラフで、特定なノード集合 V から誘導される部分グラフを次の様に定義し、以下 V 自身と同一視する。

$$G_V = \langle V, E_V \rangle (E_V = \{ \langle i, j \rangle \mid \langle i, j \rangle \in E \wedge i, j \in V \})$$

[2部グラフ] ノード集合が二つの集合に類別可能で、エッジが、この二つの集合の要素間にしか、存在しないとき、そのグラフは2部グラフであると言う。

[クリーク] 部分グラフ V が、クリークであるということは、その部分グラフが全結合の時である。

$$G_V = \langle V, E_V \rangle, \forall i, j \in V (i \neq j \rightarrow \langle i, j \rangle \in E_V)$$

[n-クリーク] V が、n-クリークであるということは、V 内の各々の要素間の G での距離 $d(G, i, j)$ が、n 以下となる場合である。

$$(\forall i, j \in V [d(G, i, j) \leq n]) \wedge (\forall i \notin V \exists j \in V [d(G, i, j) > n])$$

[n-クラン] V が n-クランであることの定義は、V が n-クリークでかつ、部分グラフ V 内での距離 $d(V, i, j)$ が n

$$V \text{ が } n\text{-クリーク} \wedge (\forall i, j \in V [d(V, i, j) \leq n])$$

下になることである。

2.2 アフィリエーションネットワーク

以下では、アフィリエーションネットワークに関連する記号について述べる。これらは主に、参考文献(4)に基づく。
[アフィリエーションネットワーク] ユーザの集団 (1~M) が、複数のグループ (M+1~M+L) に所属している時、ユーザとグループをノードとし、ユーザとグループの所属関係をエッジとするグラフを アフィリエーションネットワークと呼ぶ。

$\langle j, i \rangle, \langle i, j \rangle \in E \leftrightarrow$ 個人 $i (1 \leq i \leq M)$ がグループ $j (M < j \leq M+L)$ に所属

アフィリエーションネットワークの隣接行列は、N 人が、L グループに所属しているかどうかを表す所属行列 A を用いると、次の様に小行列を用いた形でも表現できる。

$$X_{MN} = \begin{pmatrix} O & A \\ {}^t A & O \end{pmatrix}$$

ただし、所属行列 A は、次のように定義される。

† 日本大学理工学部数学科

‡ 日本大学大学院総合科学研究科

$$A = (a_{i,j}), a_{i,j} = \begin{cases} 1 & (i \text{ 番目の個人が } j \text{ 番目のグループに所属する}) \\ 0 & (\text{その他}) \end{cases}$$

[個人間誘導ネットワーク] 個人間誘導ネットワークは、アフィリエーションネットワークから作られるグラフでノードはユーザ、エッジはユーザが共通に所属するグループがある場合に引かれる。

個人間誘導ネットワークの隣接行列も、所属行列 A を用いて、次のように求めることができる。

$$X_N = A^t A$$

[組織間誘導ネットワーク] 組織間誘導ネットワークは、組織の関係に着目したネットワークであり、ノードをグループとし、共通に所属するユーザがいる場合にエッジを引く。

組織間誘導ネットワークの隣接行列も次のようになる。

$$X_M = {}^t A A$$

2.3 中心性

個人間誘導ネットワークの分析から、組織の様々な特徴を抽出することができる。

[次数] i 番目の個人の次数は、所属行列 A を用いて次のように定義される。

$$C_D(i) = \sum_{k=1}^L a_{i,k}$$

個人間誘導ネットワークにおけるノードの次数は、「その個人がどれだけ多くの(グループを介して、他の)個人と関係をもっているか」を表現していると解釈できる。従って、その次数が大きければ大きいほど、その個人の活動量が多いことを示すことになる。

[媒介値] k 番目の個人の媒介値は、他者 ij 間の測地路中に

$$C_b(k) = \sum_{i=1}^M \sum_{j=1}^M b_{i,j}(k)$$

k が拘る ij 媒介値の総和として表される。

ただし、 ij 媒介値は、個人間誘導ネットワークの隣接行の対角要素を 0 とし、更に、個々の重みを 1 に直した行列 Y と、更に、その Y の k 行目、 k 列目の要素 0 とした行

$$Y^* = (y_{i,j}), Z^* = (z_{i,j}) \text{ の時、} b_{i,j}(k) = \frac{y_{i,j} - z_{i,j}}{y_{i,j}}$$

列 Z 、そして ij 間の最短距離 s を利用して次のように計算できる。

[フロー値] フロー値は、媒介値と同様に計算されるが、計算の途中 Y を求める時に、重みを 1 に直さず、そのまま、元の隣接行列の値を利用する。

媒介値が、媒介の有無に着目していることに対し、フロー値は媒介量にも目を注いでいる尺度になっている。

3. 組織内情報分析法

ネット上の組織活動状態は、基本的にはユーザー個々の活動で決定されるが、各ユーザーは、組織に所属することにより様々な活動や発想の可能性が与えられると共に、逆に、行為に対する制約を受けることもある。また同時に、組織内外の個人の相互作用を通して組織と個人の両方を変えてゆく可能性を持つ。複数の組織に参加する個人は、それら組織を結びつけ、その活動を通じて自身の中心性(組

織内における位置)を上げてゆくことができる。一方、高い中心性を有する個人により構成される組織は、さらに組織的な中心性を一層高めることが可能となる。このように組織活動を考えると、個人の活動だけでなく、個人から構成される組織も同時にモデル化しなければ、正確な組織内の活動分析は困難であることが分かる。そのためには、ユーザー間の接続と同時に、ユーザーが所属するグループの活動状態を同時に表現する「アフィリエーション・ネットワーク」が有効である⁽⁴⁾。グラフ理論を使い、そのネットワークを表現する場合、個人を上段、下段はグループを示すため、2部グラフ(bipartite graph)と呼ばれる。対象となる組織内活動状況は、次数中心性、媒介中心性、及びフロー中心性に関する分析結果より、リーダー的な人物、活動が盛んなグループ等を具体的に把握できる。

しかし、2部グラフは、その構成の性格上、個人間のリンクは、必ずグループを経由する必要があり、組織内に存在するグループを明確に把握する必要がある。一方、組織内グループは、名目上の組織構成とは異なり、業務やプロジェクトに応じてダイナミックに変化するものであるため、その実態を把握すること自体が課題となる。本稿では、実態に即したグループを抽出する狙いから、クリーク(対象ノード内の全てがリンクしあっている部分集合)及びクラン(半径 N のクリークで、かつ距離が N 以下となる部分グラフ)を単位とする仮想グループをメールデータから抽出する方法を実施した。

4. ケース・スタディ

4.1 組織概要と評価環境

総勢 79 名から構成され、3つの研究開発グループと支援グループ、及び事務局からなる組織を対象に、組織内情報分析フローの有効性を調査した。組織全体の運営は、研究開発運営委員会で決定され、組織全体の責任者は、各研究開発チームリーダーの持ち回りで運営される。またアシスタントチームのメンバーは、必要に応じて、研究開発チームに召集され、特定のラインには帰属しない。組織活動分析には、実際に運用されているメールサーバーに蓄積されているメール情報を使用した。今回の調査では、個人情報保護を考慮し、メールヘッダのみを対象とした。予想以上にサーバーにはスパムメールが送られてきており、その駆除に始まり、さらに、同一人物が使用している複数のメールアドレスの特定と該当するメールの整理を実施した。対象メール測定期間は 2 ヶ月で、分析ツールとしては、Mathemica⁽⁵⁾ と、NetMiner⁽⁶⁾ を使用し、データのフォーマット整理及び変換ソフトは内製した。

4.2 中心性の分析結果

(1) 実効的なグループの抽出

対象組織のクリークとクランを分析した結果を以下に示す。なお最大距離 $N=2$ とした。

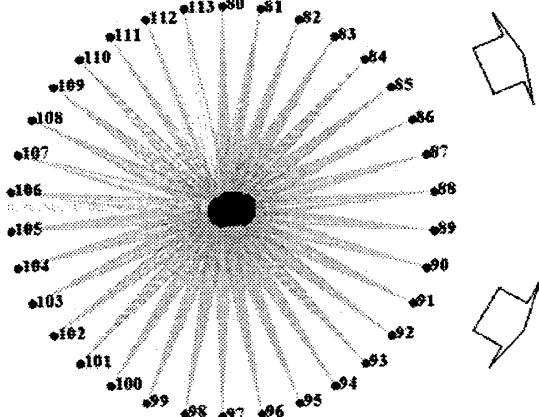
全部で 140 個のクリークが観測され、最大の凝集指数は、22.8 を示す研究員 1 名、研究支援者 2 名のコミュニティであった。凝集指数 5 以上のコミュニティは 29 グループ存在し、3 名から 6 名の小さなコミュニティで構成されている。総務のメンバーが含まれることが多いが、組織分類通りの研究員と支援者だけのグループと、組織横断的に接続を持ち活動しているメンバーにより構成されるバーチャルグループが存在することが分かった。次に、

コミュニティ内の結びつきを強め、より組織活動に近い分析を行うために、クラン分析を行った。最大 50 名の参加者があるコミュニティから、最小 6 名の参加者のいるコミュニティまで 34 個のクランが構成されていることが分かった。凝集指数は 1.50 から 4.23 の間となった。

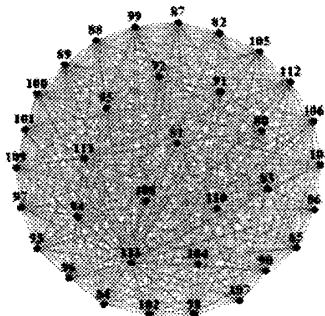
(2)アフィリエーションネットワークの生成

クラン分析結果で得られたグループ構成とメンバーデータを基に、アフィリエーション・ネットワークを表現するための 2 部グラフを形成した。ネットワークは、個人とグループを各要素とする 113 x 113 の隣接行列 (B) から描写できるが、その行列の転置行列 (tB) と B との積から、対角成分を 0 にする作業により、個人間誘導ネットワークと組織間誘導ネットワークを導くことができる。前者は、各個人対が共通に所属するグループの数を表現し、後者は、特定の組織対に共通に所属する個人の数を表したものとなる。これらの行列計算結果を元に、組織重複度を関係強度とする個人ネットワークと組織間の個人重複度を関係強度とする組織間ネットワークを導き出すことが出来、そのデータから、組織内の活動状況を具体的に把握することが出来る。組織間誘導ネットワークは、34 ノード間に完全なメッシュ上に 561 本のリンクが張られ、完全グラフとなった。また個人間誘導ネットワークは、79 ノード間に 310 本のリンクが張られる。図 1 に、得られたアフィリエーション及び誘導ネットワークの各図を示す。

個人-組織の
アフィリエーションネットワーク
(外周にグループ、中心部に個人が集合)



組織間誘導ネットワーク



個人間誘導ネットワーク

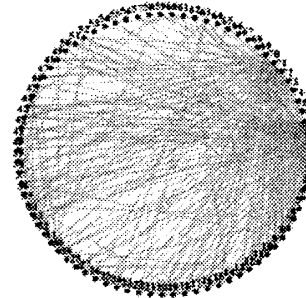


図1 アフィリエーションネットワークと誘導ネットワーク

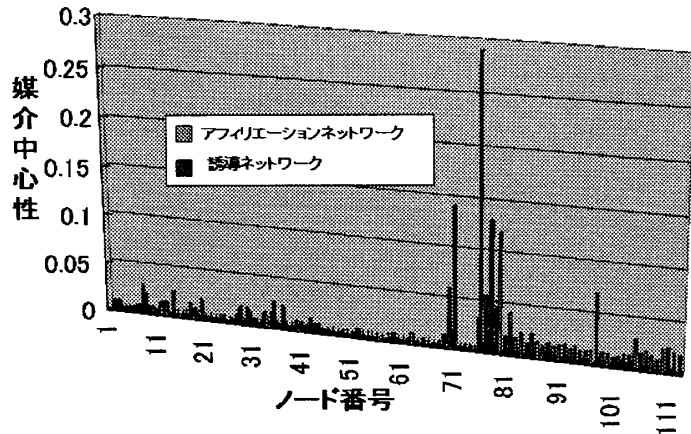


図2 媒介中心性分析結果

(3)分析結果

次数・媒介・フローの各中心性を、アフィリエーションネットワーク (モデル1) と、誘導ネットワーク (モデル2) に分けて分析した。参考として、媒介中心性に関する分析結果を図2に示す。

次数中心性に関しては、総務系ノード69~79を除くと、ノード8が、モデル1、2共に最大の次数を示し、ノード18, 30, 35, 43, 71などの次数が高いことが分かった。また媒介中心性に関しては、次数中心性と同様に、ノード8が高い媒介中心性を示す。また組織的な活動としては、次数の高いノード80が、媒介中心性も高いが、ノード98のように、次数中心性は低いにも拘らず、媒介中心性が高いノードが発見できる。つまり、組織内において、他のグループとの繋がりは相対的に低い、そのグループの成果が組織全体の活動に与える影響が大きい位置に存在していることが分かる。フロー中心性に関しては、

ノード18と28、35は、組織全体的なネットワークでの活動よりも、個人的なネットワーク内での活動の方が盛んであり、また逆に、ノード7, 10, 11は、個人的なネットワークでの中心性は低くても、組織全体のネットワークでの活動が大きい。後者の3ノードが、他のグループへ接続することによって中心性を得たことを意味する。また、前記の2中心性でトップであったノード8は、個人ネットワーク、組織全体のネットワークにおいても高い中心性を示すが、ノード18より個人ネット

