

IBM SPSS Modeler Social
Network Analysis 15 ユーザー
ガイド



注：この情報とサポートされている製品をご使用になる前に、「注意事項」（p.26）の一般情報をお読みください。

本版は IBM SPSS Modeler Social Network Analysis 15 , および新版で指示されるまで後続するすべてのリリースおよび変更に対して適用されます。

Adobe 製品のスクリーンショットは Adobe Systems Incorporated の許可を得て転載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。

Licensed Materials - Property of IBM

© Copyright IBM Corporation 2012.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

はじめに

IBM® SPSS® Modeler Social Network Analysis は、人々の関係に関する情報を、ソーシャル ネットワークにおける個人の役割を説明するフィールドに処理し、社会情報を予測モデルに含めることができますようにします。このマニュアルでは、ストリームにノードを追加できる、IBM® SPSS® Modeler 環境での IBM SPSS Modeler Social Network Analysis ノードの使用方法について説明します。ノード出力を個人の指標を示すフィールドを組み合わせると、個人のより詳細なプロフィールが作成されます。

IBM Business Analytics について

IBM Business Analytics ソフトウェアは、意思決定者がビジネス パフォーマンスを向上させるために信頼する完全で、一貫した正確な情報を提供します。ビジネス インテリジェンス、予測分析、財務実績および戦略管理、および分析アプリケーションの包括的なポートフォリオを利用することによって、現在の実績を明確、迅速に理解し、将来の結果を予測することができます。豊富な業界のソリューション、実績ある実例、専門サービスと組み合わせ、さまざまな規模の組織が、高い生産性を実現、意思決定を自信を持って自動化し、より良い決定をもたらします。

このポートフォリオの一部として、IBM SPSS Predictive Analytics ソフトウェアを使用する組織は、将来のイベントを予測し、その洞察に基づいて積極的に行動し、より優れた業績を実現することができます。全世界の企業、政府、学術分野のお客様が IBM SPSS の技術を活用し、不正行為を減少させ、リスクを軽減させながら、顧客の獲得、保持、成長において、競争優位を高めることができます。IBM SPSS ソフトウェアを日々の業務に取り入れることによって、組織は業務目標を達成し、大きな競争的優位を獲得することができますよう、意思決定を方向付け、自動化することができますようになります。お問い合わせは、<http://www.ibm.com/spss> を参照してください。

テクニカル サポート

お客様はテクニカル サポートをご利用いただけます。IBM Corp. 製品の使用方法、または対応するハードウェア環境へのインストールについてサポートが必要な場合は、テクニカル サポートにご連絡ください。テクニカル サポートの詳細は、IBM Corp. Web ページ <http://www.ibm.com/support> を参照してください。ご本人、組織、サポートの同意を確認できるものをご用意ください。

内容

1	IBM SPSS Modeler Social Network Analysis	1
	IBM SPSS Modeler Social Network Analysis について	1
	ソーシャル ネットワーク分析について	2
	ネットワークの表示	3
	ネットワークの説明	4
	IBM SPSS Modeler Social Network Analysis ノード	7
	ノードのタブ	7
	出力のプレビュー	8
	データの分析	8
	アプリケーション	9
	データ構造	9
	スクリプトのプロパティ	10
2	グループ分析	12
	グループ分析の概要	12
	社会的類似性の指定	12
	グループへの区分	13
	グループおよびグループ メンバーの説明	14
	グループ分析ノード	15
	グループ分析のデータの指定	15
	グループ分析のビルド オプションの設定	16
	グループ分析統計の表示	17
	グループ分析の出力	18
3	拡散分析	20
	拡散分析の概要	20
	拡散分析の例	21
	拡散分析ノード	22
	拡散分析のデータの指定	22
	拡散分析のビルド オプションの設定	23
	拡散分析統計の表示	24
	拡散分析の出力	24

付録

A 注意事項	26
用語集	28
参考文献	29
索引	30

IBM SPSS Modeler Social Network Analysis

IBM SPSS Modeler Social Network Analysis について

モデリングに対する多くの手法では、個人に重点を置いています。それらの手法では個人に関するさまざまなデータを使用し、行動に関する主要な指標を使用するモデルを生成します。個人に動作の発生に関連する主要指標の値がある場合、その個人に対し、その行動を回避するために特別な注意を払うことができます。

顧客が企業との関係を断ち切る、解約のモデリングに対する手法について考えます。顧客を維持するコストは、顧客を入れ替えより大幅に低く、解約のリスクにある顧客を特定することができます。アナリストは、人口統計情報、各顧客の最近の通話パターンなど、多くの主要業績評価指標を使用し、顧客に関する説明を行います。こうしたフィールドに基づいた予測モデルでは、過去に解約した顧客の通話パターンと一貫性のある顧客の通話パターンの変化を活用して、解約のリスクが高まっている顧客を特定します。リスクありと判断された顧客は、彼らを慰留するため、追加のカスタマー サービスまたはサービス オプションを受けます。

これらの方法では、顧客の行動に大きく影響を与えるとされる社会的情報を見落としします。企業に関する情報、そして顧客以外の人々の行動に関する情報は関係を超えて拡大し、人々に影響を与えます。その結果、他の人々との関係によって、彼らがある人の意思決定や行動に影響を与えます。個人の指標の実を使用する分析では、よそき機能を持つ重要な要素を除外してしまいます。

IBM® SPSS® Modeler Social Network Analysis では、関係に関する情報をモデルで使用できるフィールドに処理することによって、この問題に対応します。派生したこれらの主要業績評価指標は、個人の社会的特性を評価します。こうした社会的プロパティを個人ベースの指標を組み合わせることによって、個人に関する概要がより詳細になり、その結果モデルの予測の精度が向上します。

IBM SPSS Modeler Social Network Analysis は、次のコンポーネントから構成されています。

- IBM® SPSS® Modeler 環境に追加された IBM SPSS Modeler Social Network Analysis ノード。ストリームでの社会的分析手法を活用できます。
- ノードの特性の処理を IBM® SPSS® Modeler Server に追加する IBM® SPSS® Modeler Server Social Network Analysis。IBM SPSS Modeler Server Social Network Analysis は膨大な量のネットワークデータを効果的に処理し、多くの個人および関係を比較的少ない数のフィールドに投入してより詳細に分析することができるようになります。

たとえば IBM SPSS Modeler Social Network Analysis は、解約する特定の人々に最も影響を受けるネットワーク内の個人を特定します。さらに、解約のリスクが高いネットワーク内の個人のグループを見出すことができます。モデルにこれらの影響の主要業績指標を組み込むことによって、パフォーマンス全体を向上させることができます。

注： IBM SPSS Modeler Social Network Analysis ノードの処理を、IBM SPSS Modeler Server Social Network Analysis によって行う必要があります。SPSS Modeler で IBM SPSS Modeler Social Network Analysis ノードを含むストリームを実行する場合、IBM SPSS Modeler Server Social Network Analysis も含む SPSS Modeler Server にインスタンスに接続する必要があります。

ソーシャル ネットワーク分析について

ソーシャル ネットワークは複数の個人とその個人間の関係で構成されています。ソーシャル ネットワーク分析では、これらの関係を検証し、個人およびグループを社会構造の一部として説明します。個人はお互いに相互作用し、これらの相互作用のパターンにより、関係する個人について理解を深めることができます。関係によって情報をネットワーク全体に広めることができ、個人が他人に対して影響を与えることができます。関係情報の重要度が、ソーシャル ネットワーク分析が他の手法と異なる点です。各個人それぞれに重点を置く代わりに、調査の単位は 2 人の個人および彼らの間の関係で構成される対となります。

ネットワークの関係は、方向付き関係と方向なし関係に分類されます。**方向付き関係**では、一方の個人が関係の主導者、またはソースとして、もう一方の個人が受動者またはターゲットとして識別されます。たとえば、電話をかけることは、ある人が他社にかける方向付き関係となります。それに対し、**方向なし関係**の場合、ソースおよび宛先の役割を定義することはできません。この場合、二者とも関係においては同等となります。方向なし関係の例として、お互いに会話することがあげられます。

関係を区別するもう 1 つのプロパティが、関係が二分であるか評価付きであるかです。**二分関係**で使用できる情報は、関係が 2 個人間にあるかどうかだけです。ネットワークのすべての対について、関係が存在する場

合と存在しない場合があります。それに対し、**評価付き関係**には、関係の強度を示す重みを使用します。重みを使用することによって、関係を互いに比較することができます。

“**関係の種類**” 表に、方向およびスケールによる交差分類の例を示します。方向付き関係においては、Joe が関係のソース、Mary がターゲットとなります。方向なし関係においては、誰が関係の基点となるかの識別はありません。評価付き関係では、関係の重みとして会話の長さを利用しますが、二分関係では会話があったかどうかを利用します。

テーブル 1-1
関係の種類

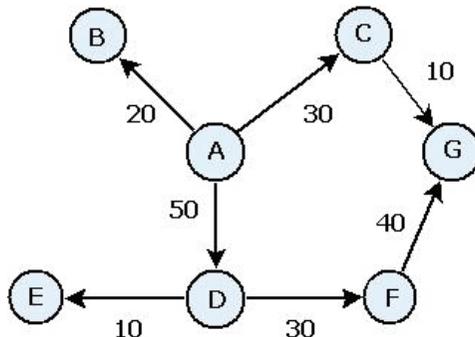
方向	スケール	例
方向なし	二分	Joe と Mary が会話する
方向なし	評価付き	Joe と Mary が 20 分間会話する
方向あり	二分	Joe が Mary に電話する
方向あり	評価付き	Joe が Mary に電話して 20 分間会話する

ソーシャル ネットワーク分析のフィールドの詳細は、この分野に関する書籍 (Wasserman および Faust, 1994) のいずれかを参照してください。

ネットワークの表示

ソーシャル ネットワークは通常、**ソシオグラム** (Moreno, 1934) を使用して表します。こうした視覚的な表示では、個人は空間における点またはノードに対応します。点を結ぶ線または辺は個人間の関係を示します。関係に方向がある場合、辺はその方向を示す矢印となります。関係に重みがある場合、辺のラベルが値を示します。次のグラフは、7 人のネットワークを示します。

図 1-1
ソーシャル ネットワークの例



ネットワークが電話による通話を示し、関係の重みが通話の長さを示すとします。この場合、個人 A が 3 人に電話し、D との会話に多くの時間を費やしていることを示します。

このネットワークは実際のものより小さいものですが、このシンプルなソシオグラムが示す概念は、サイズおよび複雑さに関係なくネットワークを一般化したものです。

ネットワークの説明

ネットワーク、グループ、個人のに関する情報を、お互いに比較して予測モデルに利用することができる説明的特性に抽出する必要があります。ネットワークは、限定された主要業績評価指標に抽出して分析できるようにする必要があります。たとえば、ネットワーク内のノードまたはノードのグループをお互いに比較する場合があります。または、ネットワーク内の個人をお互いに比較したり、最も重要な個人を特定する場合があります。

ソーシャル ネットワークを説明するために使用する代表的な指標として**密度**と**程度**があります。どちらの統計情報も関係性を反映しますが、密度はネットワーク全体またはネットワーク サブグループに重点を置き、程度はネットワークの個人を特徴づけます。

ネットワークの密度

ネットワークのノードについて、使用できる関係の数は限られています。それぞれのノードが、他のノードすべてとの関係のソースまたはターゲットとして機能します。3 つのノード A、B、C で構成されたネットワークがあるとします。次の表は、ノード間の方向付き関係を示します。

テーブル 1-2
3 つのノードの方向付き関係

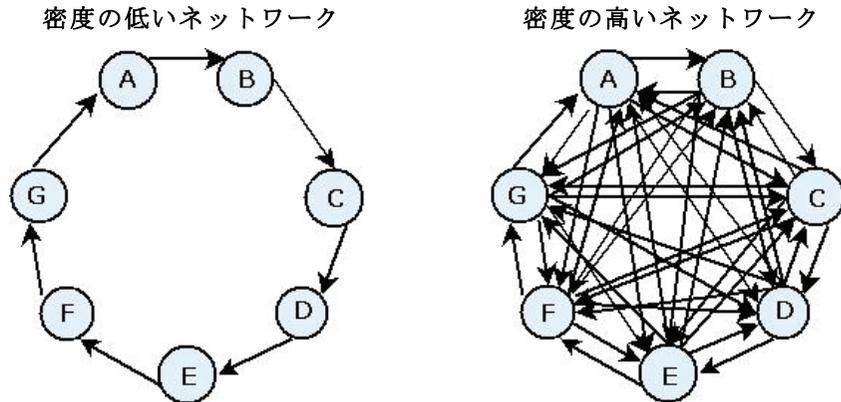
ソース	ターゲット
A	B
A	C
B	A
B	C
C	A
C	B

各ノードが他の 2 つのノードとの関係のソースとなります。ただし実際は、考えられるすべての関係が存在するわけではありません。ノードが他のノードと直接関係を持たない場合もあります。また、方向付き関係が相互一致しない場合もあります。

密度の統計情報は、実際に存在するネットワークの関係の割り合いを示します。値の範囲は 0 ～ 1 です。0 は関係が存在しないネットワーク、1 は考えられるすべての関係が存在するネットワークを示します。値が 1 に近いほど、ネットワークの密度が高くなり、ネットワーク内のノードが密接します。

密度が高いネットワークの情報は、希薄なネットワークと比較して拡散しやすくなります。“密度の低いネットワークと高いネットワーク”表は、7 つのノードで構成された 2 つのネットワークを示します。密度が低いネットワークにはノード間で考えられる 42 の関係のうち、7 つの関係だけが存在し、密度が 0.17 となります。密度の高いネットワークでは、考えられるすべての関係が存在し、密度が 1.0 となります。

テーブル 1-3
密度の低いネットワークと高いネットワーク



密度の低いネットワークの場合、情報をノード A からノード G に拡散するには、他の 5 つのノードを通過する必要があります。それに対し、密度の高いネットワークでは、情報をノード A から G に直接送ることができます。

ノードの度数

ネットワーク内で重要な個人は、多くの関係に関わっています。重要な個人はさまざまなソースから情報を取得し、その情報を多くの個人に拡散します。それに対し、関係が少ない個人は、ネットワーク内の多くの人々に直接影響を与えることはできません。

そのノードが関わる関係数の合計として定義されるノードの**度数**によって、ネットワーク内の個人間で比較することができます。度数が高い個人は、度数が低い個人と比較してアクティブになります。度数は、関係の方向を無視し、ノードの活動を示す全体的な指標を提供します。

方向付き関係では、関係の数を集計する場合、ノードがソースなのかターゲットなのかを重要視します。ノードの **Indegree** は、特定のノードがターゲットとなる関係の数を示します。それに対し、**Outdegree** はノードがソースとなる関係の数を示します。次のテーブルは “ソーシャル ネットワークの例” 図の各ノードの度数、Indegree、および Outdegree の値を示します。

テーブル 1-4
度数、Indegree、Outdegree 値の例

ノード	度数	Indegree	Outdegree
A	3	0	3
B	1	1	0
C	2	1	1
D	3	1	2
E	1	1	0
F	2	1	1
G	2	2	0

Indegree は、地位の指標として扱われる場合があります。Indegree の値が高いほど、多くの関係がそのノードをターゲットとします。つまり、多くの人から接触されていることを示します。他の多くのノードがそのノードとの関係の起点となっています。それに対し、Outdegree は、中心性の指標として扱われる場合があります。Outdegree の値が高いほど、多くの関係がそのノードをソースとします。こうした個人から、他の多くの人々に接触しています。

ネットワーク例でのノードについては、度数の値はノード A および D が最もアクティブで、ノード B と E がアクティブでないことを示しています。Indegree 値は、ノード G の地位が最も高いことを示しています。Outdegree 値に基づくと、ノード A が中心となります。

IBM SPSS Modeler Social Network Analysis ノード

IBM® SPSS® Modeler で提供される多くの標準ノードとともに、IBM® SPSS® Modeler Social Network Analysis ノードと連携して、ストリームでのソーシャル ネットワーク分析の結果を利用することもできます。“IBM SPSS Modeler Social Network Analysis ノード” 表は、[ソース] パレットに格納されているこれらのノードについて説明します。

テーブル 1-5
IBM SPSS Modeler Social Network Analysis ノード

ノード	アイコン	説明
グループ分析		グループ分析ノードは、固定フィールド テキストファイルから通話の詳細レコード データをインポートし、レコードで定義されたネットワーク内のノードのグループを識別、ネットワークのグループや個人の主要業績評価使用を生成します。詳細は、2 章 p.12 グループ分析の概要 を参照してください。
拡散分析		拡散分析ノードは、固定フィールド テキストファイルから通話の詳細レコード データをインポートし、レコードで定義されたネットワーク間の効果を拡張、個人のノードに対する効果の結果を要約する主要業績評価使用を生成します。詳細は、3 章 p.20 拡散分析の概要 を参照してください。

IBM SPSS Modeler Social Network Analysis ノードの処理を、IBM® SPSS® Modeler Server Social Network Analysis によって行う必要があります。SPSS Modeler でこれらのノードを含むストリームを実行する場合、IBM SPSS Modeler Server Social Network Analysis も含む IBM® SPSS® Modeler Server にインスタンスに接続する必要があります。

ノードのタブ

IBM® SPSS® Modeler Social Network Analysis ノードには、以下のように分析を定義およびプレビューするためのタブがあります。

- **[データ] タブ:** ソーシャル ネットワーク情報を含むファイルを指定します。
- **[ビルド オプション] タブ:** 分析の設定を定義します。
- **[分析] タブ:** [データ] タブの入力設定を変更して最適な結果を生成するためのガイダンスを提供する、準備用出力の要約を表示します。

また、ノードには IBM® SPSS® Modeler のソース ノードで共通する以下のようなタブが表示されます。

- **[フィルタ] タブ:** ノードで作成された出力フィールドを除外したり名前を変更したりします。このタブは、フィルタ ノードと同じ機能を提供します。
- **[データ型] タブ:** ノードで作成された出力フィールドの尺度を設定します。このタブは、データ型ノードと同じ機能を提供します。
- **[注釈] タブ:** 名前の変更、ユーザー指定のツールヒント、長い注釈の格納を行います。

共通タブの詳細については、SPSS Modeler のマニュアルを参照してください。

出力のプレビュー

通常、ノードの出力は予測モデルへの入力として使用します。出力を表示する場合、テーブル ノードをストリームに追加しますが、データの量が非常に多く、テーブルの生成に時間がかかります。また、テーブルの行数によって、使いやすさが限られます。

作成される出力のサンプルを表示するプレビュー テーブルを生成することができます。プレビューには、一部の行の生成フィールドが表示されます。行数は、ストリーム プロパティで指定します。ノード出力をプレビューするには、次の手順に従います。

1. ノードを開きます。
2. [データ] タブでデータ設定を指定します。
3. [ビルド オプション] タブで分析パラメータを指定します。
4. [プレビュー] をクリックします。

プレビュー ウィンドウが開きます。また、出力をプレビューすると、ノードの [分析] タブに結果の概要が表示されます。

データの分析

最適な結果を生成する分析設定を指定するには、通常繰り返し設定作業を行います。設定を定義、分析を実行、結果をレビューします。結果が薬に他立たない場合、設定を変更して分析を再実行します。

ノードの入力データを分析するには、次の手順に従います。

1. ノードを開きます。
2. [データ] タブでデータ設定を指定します。
3. [ビルド オプション] タブで分析パラメータを指定します。
4. [データの分析] をクリックします。

ビルド オプションで要約統計を表示するよう設定した場合、[分析] タブに結果が表示されます。

分析を再実行する必要がある場合、[分析をクリア] をクリックし、現在の結果を消去してから [データの分析] をクリックします。

アプリケーション

IBM® SPSS® Modeler Social Network Analysis が特に役に立つアプリケーションには以下のようなものがあります。

- **解約予測:** グループの特性が解約率に影響を与える場合があります。解約のリスクが高いグループの個人に重点を置くことによって、それを回避することができます。また、すでに解約した人物からの情報の拡散によって解約のリスクがある個人を特定できます。
- **グループリーダーの利用:** グループ リーダーは、他のグループ メンバーに対し影響力が大きくなります。グループ リーダーの解約を回避した場合、グループ メンバーの解約率を軽減することができます。または、競合他社のグループ リーダーを解約させることにより、その競合他社に関係するグループ メンバーの解約率を拡大し、自社に関係するグループ メンバーの解約率を軽減することができます。
- **マーケティング:** グループ リーダーを使用して、新しい商品またはサービスを広めることができます。リーダーの影響により、他のグループ メンバーがサービスを購入する確率が高まる場合があります。拡散分析を使用してグループ リーダーの影響を最も受ける個人を特定し、マーケティングの対象を彼らに設定します。

データ構造

関心のあるネットワークの個人に関する情報は、企業内のさまざまなファイル、データベースおよびシステムで拡散させることができます。IBM® SPSS® Modeler Social Network Analysis を使用してネットワークを分析するには、関連するレコードおよびフィールドをデータ ソースから抽出し、ノードの入力できるようにそれらの形式を設定します。

分析ノードでは、1 つの固定幅テキスト ファイルの保存されている通話詳細レコードが必要です。ファイルの各行は関係に対応しており、データは以下の列に整理されています。

- 関係の起点となる個人の識別子
- 関係のターゲットとなる個人の識別子
- 関係の重み（オプション）

すべてのデータは数値でなければなりません。個人の識別子は整数です。オプションで、ファイル名をデータの 1 行目に指定します。“通話データ レコードの例” 表は、このデータの構造を示します。

テーブル 1-6
通話データレコードの例

ソース	ターゲット	重み
1000	5642	243
2190	8444	831
0299	9419	559

重みの値は、ネットワークのほかの関係に対する関係の重要度を示す場合に使用する指標に対応させることがあります。通話データの場合、一般的な重みには通話時間や通話の頻度があります。

分析を通話履歴のサブセットに重点を置く場合、入力ファイル作成時にそのサブセットを使用する必要があります。たとえば、入力テキスト ファイルに過去数ヶ月間または最近の通話のデータを含めることによって、分析をこれらのデータに制限することができます。

スクリプトのプロパティ

IBM® SPSS® Modeler Social Network Analysis のスクリプトは、ユーザーインターフェースのプロセスを自動化する強力なツールです。スクリプトで、マウスやキーボードを使用した場合と同じ種類のアクションを実行できます。また、頻繁に繰り返したり手動で実行するのに時間がかかるタスクを自動化するために使用できます。スクリプトの使用例については、IBM® SPSS® Modeler に付属の ScriptingAutomation.pdf ガイドを参照してください。

拡散ノードのプロパティ

次の表は、拡散ノードのスクリプト プロパティを示しています。

プロパティ名	データ型	プロパティの説明
input_data_file_name	string	
calling_field	フィールド	
called_field	フィールド	
frequency_weight_field	フィールド	
read_field_names	boolean	
diffusion_list_file_name	string	
spreading_factor	double	デフォルト値 = 50 最小値 = 1 最大値 = 99
max_number_iterations	integer	デフォルト値 = 100 最小値 = 1

プロパティ名	データ型	プロパティの説明
accuracy_threshold	double	デフォルト値 = 0.01 最小値 = 0.001
calculate_statistics	boolean	

グループ分析ノードのプロパティ

次の表は、グループ分析ノードのスクリプト プロパティを示しています。

プロパティ名	データ型	プロパティの説明
input_data_file_name	string	
calling_field	フィールド	
called_field	フィールド	
frequency_weight_field	フィールド	
read_field_names	boolean	
coverage_threshold	double	デフォルト値 = 10 最小値 = 1 最大値 = 99
min_group_size	integer	デフォルト値 = 2 最小値 = 2
max_group_size	integer	デフォルト値 = 100 最小値 = 2
calculate_statistics	boolean	

グループ分析

グループ分析の概要

グループ分析では、ネットワーク内の個人の相互関係のパターンを使用し、類似した個人のグループを特定します。これらのグループの特性は、各グループメンバーの行動に影響を与えます。たとえば、メンバー間の関係が多く、強いリーダーのいる小規模なグループは、グループのメンバーが実際に解約していない場合でも、解約のリスクが高くなります。(Richter, Yom-Tov, および Slonim, 2010) グループと個人の指標を組み合わせる予測モデルは、個人の指標のみを使用するモデルより実績が高まります。

グループ分析は、次の手順で行います。

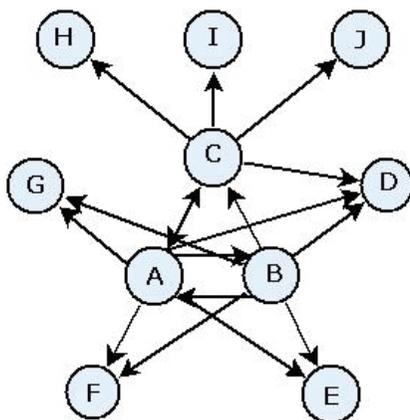
1. 社会的類似性を反映する関係の強度を指定します。詳細は、p. 12 [社会的類似性の指定](#) を参照してください。
2. サイズの制限に従いながら関係の強度に基づいてネットワークをグループに区分します。詳細は、p. 13 [グループへの区分](#) を参照してください。
3. グループリーダーを指定するほか、グループと個人のプロフィールを作成します。詳細は、p. 14 [グループおよびグループメンバーの説明](#) を参照してください。

社会的類似性の指定

グループのメンバーは、グループ外の個人と比較してより類似していなければなりません。ネットワーク分析に置いて、2つのノードの類似性はそれらの関係によって異なります。ネットワーク内のノードについて、方向が指定されたノードとの関係のターゲットとなる一連のノードがあります。電気通信データの場合、このノードのセットは、特定の個人が連絡したすべての人物に対応します。2人の人物が同じグループに連絡を取った場合、彼らはお互いに類似していると見なされます。2人の個人についてより多くの関係のターゲットのセットが重複すると、類似性がより高まります。

“10 ノード ネットワークの例” 図に示されるネットワークについて考えてみましょう。

図 2-1
10 ノード ネットワークの例



“ターゲット ノード” 表では、ノード A、B、および C を起点とする関係のターゲット ノードを示します。

テーブル 2-1
ターゲット ノード

ソース ノード	ターゲット ノード
A	B、C、D、E、F、G
B	A、C、D、E、F、G
C	A、D、H、I、J

ノード A と B には 5 つの共通ターゲット ノードがあります。それに対し、ノード A と C の共通ターゲット ノードは 1 つだけです。ノード A と B は、ノード A と C と比較して社会的類似性が高くなります。

通話時間または通話頻度のような関係の重みは、ネットワークのノードの類似性を取得するものではありません。類似性のイメージを加えるために、グループ分析では、共通の情報（Cover および Thomas, 2006）を関係の重みとして使用します。この統計は、2 つのノードが同じノードに接続するという確率を反映しています。ネットワーク内の関係に、電気通信データの通話頻度などの定義された重みがある場合、共通の情報がこれらの重みを随時組み込みます。

グループへの区分

グループは、他のグループ メンバーとの関係の重みが大きい個人で構成されます。重みは、関係ないのノードの類似性の指標となります。その結果、グループの指定では、まずネットワーク内の関係性が弱いものを除外します。**カバレッジ閾値**は、保持する最も強い関係の割合を定義してこのプロセスを管理します。たとえば、カバレッジ閾値の値が 0.4 の場合、

最も強い 40% の関係性がグループの指定に使用され、残りの 60% の関係が除外されます。

残りの関係によって、予測ユーティリティを制限した、非常に小さいグループ、または非常に大きいグループを生成する場合があります。こうしたグループが分析に含まれないようにするために、グループの最小サイズ、最大サイズを指定することができます。最小サイズより小さいグループは完全に除外されます。最大サイズより大きいグループは、可能なサイズの範囲内に分割されます。弱い関係性を削除しサイズの上限を指定した後残ったグループは、**コア グループ**と呼ばれます。

元のネットワークとの関係を削除すると、何人かの個人がコア グループに割り当てられなくなる場合があります。ただし、これらの個人はグループへの割り当てが保証されているグループのメンバーと関係がある場合があります。グループのコア メンバーに対して比較的強い関係が多くある場合、グループのサイズ制限に反していなかぎり個人が追加されます。そのため、最終的なグループは、コア メンバーおよびコアとの関連性によって追加されたメンバーで構成されます。

グループおよびグループ メンバーの説明

密度、Indegree、Outdegree のほか、その他の統計情報がグループの力学について説明します。具体的には、権限および拡散スコアがグループ内の個人のソーシャル ステータスの指標を提供します。グループとメンバーの行動を予測しようとする場合、グループ内の各個人の役割が極めて重要となります。

ノードの**権限スコア**は、ノードに関連するグループ内の他のノードの傾向を測定します。情報や意見を求めて、多くの個人が特定の個人に接触している場合、その個人には権限の役割があります。グループ ネットワークで再開するランダム ウォークの定常確率に対応する、グループのノードの権限スコアは 0 ~ 1 の範囲です。権限スコアが 1 に近いほど、グループ内のノードの権限スコアが大きくなります。権限スコアが最も大きいグループのノードは、グループの**権限リーダー**となります。グループの最大スコアを最少スコアで割ると、権限リーダーの全体の強度の指標が分かります。

一方、ノードの**権限スコア**は、グループ内の他のノードに関連するノードの傾向を評価します。特定の人物がグループ内の他の多くの人物に連絡を取る場合、その人物はグループ全体の意見の大きく影響を与える場合があります。逆方向でグループ ネットワークで再開するランダム ウォークの定常確率に対応する、グループのノードの拡散スコアは 0 ~ 1 の範囲です。拡散スコアが 1 に近いほど、他のグループ メンバーに接触するノードが多くなります。拡散スコアが最も大きいグループのノードは、グルー

プの**拡散リーダー**となります。グループの最大スコアを最少スコアで割ると、拡散リーダーの全体の強度の指標が分かります。

グループ分析ノード

[ソース] パレットから使用できるグループ分析ノードは、お互いに社会的に類似したネットワーク内の個人を特定し、グループ内の個人の相対的なソーシャル ステータスを指定します。比較的小さい規模でメンバー間の接触が多く、強力なリーダーがいるグループは、グループの力学により解約のリスクが高くなります。個人の解約スコアを組み合わせると、その結果を使用してグループの解約スコアを確認し、個人のスコアのみに基づくよりもモデルから解約を予測できるようになります。

データが通常ネットワークと、グループ定義を基底とする社会的相互作用の性質を構成する場合、グループ分析は時間がかかるプロセスとなります。通常、グループの分析結果は予測モデルの入力として使用するデータベースまたはファイルに保存します。グループの結果は月 1 回などの頻度で更新されますが、モデルはより頻繁に更新する場合があります。

要件: ノードには、3 つのフィールドを使用してソーシャル ネットワークを定義する固定幅のテキストファイルが必要です。1 つ目のフィールドは方向が指定された各関係のソースを定義、2 つ目のフィールドは方向が指定された各関係の宛先を定義、3 つ目のフィールドは各関係の任意の強度を指定します。ネットワーク内のすべての関係には方向を指定する必要があります。

グループ分析のデータの指定

[グループ分析] ソース ノード ウィンドウの [データ] タブを使用して、ネットワーク ノードの関係を指定する入力ファイルを指定できます。

ファイル: 通話詳細レコードを含むファイルまたはフォルダの名前を指定します。名前を入力するか、[...] ボタンをクリックしてファイル システムの名前を選択します。名前を選択すると、下部にファイル パスが表示されます。また、ファイルの内容が、区切り文字付きで表示されます。フォルダを指定する場合、フォルダに含まれるすべてのファイルの呼び出しの詳細レコードが連結され、分析されます。フォルダのすべてのファイルの構造は同じでなければなりません。

ファイルからフィールド名を取得: デフォルトで選択されているこのオプションは、データ ファイル中の最初の行を列の名前として取り扱います。1 行目が見出しでない場合、このオプションをオフにして各フィールドに自動的に Field1 や Field2 といった汎用名を割り当てます。

ネットワーク定義設定

ネットワーク定義設定はフィールドの役割を定義します。

フィールド: 矢印ボタンを使用して、このリストから項目をさまざまな役割フィールドに手動で割り当てます。アイコンは、各役割フィールドの有効な測定レベルを示します。[すべて] ボタンをクリックしてリスト内のすべてのフィールドを選択するか、各測定レベルのボタンをクリックして、その測定レベルのすべてのフィールドを選択します。

ソース: 方向付き関係の起点としてフィールドを 1 つ選択します。

ターゲット: 方向付き関係のターゲットとしてフィールドを 1 つ選択します。

重み: 任意で、ネットワーク内に関係が出現する回数を示すフィールドを 1 つ選択します。関係の強度を定義する場合、分析によりレコードに重みが付けられます。

グループ分析のビルド オプションの設定

[グループ分析] ソース ノード ウィンドウの [ビルド オプション] タブを使用して、ネットワークのグループを特定するためのオプションを定義できます。

グループ分析設定

グループ分析設定は、形成されたグループのサイズおよび相対的な強度に影響を与えます。

カバレッジしきい値: 分析で使用する最も強力なネットワーク関係の割合を定義します。たとえば、カバレッジしきい値が 0.2 の場合、重みが上位 20% の関係のみが使用されることを示します。範囲が 0 ~ 1 のこのパラメータを使用して、分析をグループ内の最も強力な関係に焦点を当てることができます。最も大きいしきい値は、メンバー間の関係が最も弱いグループを示します。詳細は、p.13 [グループへの区分](#) を参照してください。

最小グループ サイズ: グループ サイズの下限を指定します。この値より小さいグループは返されません。

最大グループ サイズ: グループ サイズの上限を指定します。この値より大きいグループはより小さいグループに分割されます。

要約統計量の計算および表示: このオプションを選択すると、ノードは主要業績評価指標出力を生成するほか、分析の要約統計量を計算して表示します。ネットワークが極めて大きい場合、これらの統計量の計算は、ノードのパフォーマンスに悪影響を与える場合があります。

グループ分析統計の表示

[グループ分析] ソース ノード ウィンドウの [分析] タブには、ネットワーク内で識別されるグループの概要が表示されます。“グループ分析の要約統計” 表に、利用できるグループの要約統計を示します。

テーブル 2-2
グループ分析の要約統計

統計	説明
グループ内のノード数	指定されたグループ内のノード数
グループ内のリンク数	指定されたグループ内のリンク数
グループ数	ネットワーク内のグループ数
平均グループ サイズ	グループ内の平均ノード数
平均グループ密度	グループ内のノード間の直接的接触の割合の平均 詳細は、1 章 p.4 ネットワークの密度 を参照してください。
コアメンバーの平均割合	グループのコア ノードとなるノードの割合の平均 詳細は、p.13 グループへの区分 を参照してください。
コア グループ密度の平均	グループ内のコア ノード間の直接的接触の割合の平均
平均 InDegree	受信リンクの平均数。詳細は、1 章 p.5 ノードの度数 を参照してください。
平均 OutDegree	送信リンクの平均数。詳細は、1 章 p.5 ノードの度数 を参照してください。

要約統計表から特定の統計を選択して、値の分布、標準偏差、統計の歪度を表示します。

これらの統計を確認して、グループ分析設定を指定できます。結果で密度が高い比較的小さなグループを示すのが理想的です。たとえば、非常に大きいグループがある場合、最大グループ サイズを小さくしてより小さいグループに分割してください。また、グループの密度が低い場合、カバレッジしきい値を小さくして強力な関係に重点を置きます。

グループ分析の出力

グループ分析ノードは、グループおよびグループ内の個人について説明するさまざまなフィールドを生成します。既存のモデルとデータを主要行政指標によって増補し、これらのモデルによって生成された予測を改善することができます。たとえば、個人の解約の傾向を更新してグループの影響を含めることができます。

“グループの主要業績指標” 表に、分析におけるグループの主要業績指標フィールドを示します。

テーブル 2-3
グループの主要業績指標

フィールド	説明
GAG_GroupNumber	グループの一意の識別子
GAG_Size	グループ内の人数
GAG_Density	グループ内の個人間の直接的接触の割合の平均 詳細は、1 章 p.4 ネットワークの密度を参照してください。
GAG_KernelDensity	グループ内のコア個人間の直接的接触の割合の平均
GAG_CoreNodesFraction	グループのコア個人となる個人の割合の平均 詳細は、p.13 グループへの区分を参照してください。
GAG_MaxRankType1	グループ メンバーの最大権限スコア。 詳細は、p.14 グループおよびグループ メンバーの説明を参照してください。
GAG_MinRankType1	グループ メンバーの最小権限スコア。
GAG_MaxMinRankRatioType1	最大権限スコアの最小権限スコアに対する比率。この値はグループ リーダーの権限の強度を反映します。
GAG_MaxRankType2	グループ メンバーの最大拡散スコア。 詳細は、p.14 グループおよびグループ メンバーの説明を参照してください。
GAG_MinRankType2	グループ メンバーの最小拡散スコア。
GAG_MaxMinRankRatioType2	最大権限スコアの最小拡散スコアに対する比率。この値はグループ リーダーの拡散の強度を反映します。

“個人の主要業績指標” 表に、ネットワークにおける個人の主要業績指標フィールドを示します。

テーブル 2-4
個人の主要業績指標

フィールド	説明
GAI_NodeNumber	個人の一意的識別子
GAI_CoreNode	個人がグループのコア個人がどうかの指標。 詳細は、p.13 グループへの区分を参照してください。

フィールド	説明
GAI_RankType1	個人の権限スコア。詳細は、 p.14 グループおよびグループメンバーの説明 を参照してください。
GAI_RankOrderType1	権限スコアに基づいたグループ内の順位。
GAI_RankType2	個人の拡散スコア。詳細は、 p.14 グループおよびグループメンバーの説明 を参照してください。
GAI_RankOrderType2	拡散スコアに基づいたグループ内の順位。
GAI_InDegree	個人が関係のターゲットである関係の数。詳細は、 1 章 p.5 ノードの度数 を参照してください。
GAI_OutDegree	個人が関係のソースである関係の数。詳細は、 1 章 p.5 ノードの度数 を参照してください。
GAI_GroupLeaderType1	ノードが権限リーダーかどうか。リーダーシップスコアは受信リンクから派生。詳細は、 p.14 グループおよびグループメンバーの説明 を参照してください。
GAI_GroupLeaderConfidenceType1	ノードが権限リーダーである確信度。
GAI_GroupLeaderType2	ノードが派生した拡散リーダーかどうか。リーダーシップスコアは送信リンクから派生。詳細は、 p.14 グループおよびグループメンバーの説明 を参照してください。
GAI_GroupLeaderConfidenceType2	ノードが拡散リーダーである確信度。

拡散分析

拡散分析の概要

拡散分析では、ソーシャル ネットワーク内で他人に最も影響を受ける個人を特定し、効果を**拡散エネルギー**として定量化します。このプロセスでは、効果がネットワーク ノードから近隣に繰り返し広がる活性化拡散手法を使用し、ノードからノード (Dasgupta, Singh, Viswanathan, Chakraborty, Mukherjea, Nanavati, および Joshi, 2008) に移動するごとにサイズを縮小します。エネルギーを受信すると、ノードはアクティブ化され、そのエネルギーの一部を方向が指定されたノードとの関係のターゲットとなっている近隣に送信します。

拡散係数は、一部がアクティブ化されたノードによって送信され、残りの部分がノードによって保持されるエネルギーを定義します。このエネルギーを受信するノードは、同じ割合を近隣に送信し、エネルギー送信の崩壊プロセスをもたらします。拡散係数が大きくなるとより多くのエネルギーが送信され、プロセスが完全に崩壊する前にエネルギーを最初のアクティブ化ノードからより離れたノードに送信することができます。拡散係数が小さい場合、拡散プロセスが速く崩壊し、送信されるエネルギーが初期のノードに比較的近くに残ります。

アクティブ化ノードによって拡散されるすべてのエネルギーは、方向が指定されたノードとの関係のターゲットとなるすべてのノード間に分布します。各ノードが受信するエネルギー量は、アクティブ化ノードとの関係の強度によって異なります。特定のノードに送信される割合は、関係の重みをアクティブ化ノードがソースとなるすべての関係の重みの合計で割った値となります。一方、相対的に重みの大きい関係を持つ近隣は、相対的に重みの小さい近隣より大きなエネルギーを受信します。

次の条件のいずれかを満たす場合、拡散プロセスは停止します。

- アクティブ化ノードが方向が指定された関係のソースではない
- 送信されるエネルギーの量が、プロセスを続行するための送信エネルギー量の上限である、**精度しきい値**を下回っている
- 反復数が指定の上限に達している

拡散が完了した場合、エネルギーが最も拡散しているノードがプロセスを開始した効果に対して最も感度が高くなります。たとえば、チャーン (Churn) が発生するノードで開始する場合、最もエネルギーの大きいノード

ドが最もチェーンのリスクが高くなります。チェーンが発生しないよう、これらのノードに特別の注意を払うことができます。

拡散分析の例

1 章 p.3 で示すネットワークについて考えてみます。ネットワークは、お互いに強度の異なる方向が指定された関係の 7 つのノードで構成されています。

“拡散プロセスの例” 表は、0.80 の拡散係数を使用して、ネットワーク間の拡散プロセスを示しています。最初、ノード A にはすべてのエネルギーが含まれ、任意で 1.00 の値が割り当てられています。ステップ 1 でこのノードがアクティブ化され、エネルギーの 80% が関係のターゲットとなる 3 つの近隣に拡散し、20% のエネルギーが留まります。関係の重みは合計 100 となります。ノード D との関係はこの合計の半分を占めるため、このノードは拡散エネルギーの半分または 0.40 を受信します。ノード B との関係は全重みのうち 20% を占め、0.16 の拡散エネルギー値を受信します。ノード C は残りのエネルギー 0.24 を受信、ノード A 画家発生するエネルギー量の 30% となります。

テーブル 3-1
拡散プロセスの例

ステップ	ノード A	ノード B	ノード C	ノード D	ノード E	ノード F	ノード G
0	1.00	0	0	0	0	0	0
1	0.20	0.16	0.24	0.40	0	0	0
2	0.20	0.16	0.05	0.08	0.08	0.24	0.19
3	0.20	0.16	0.05	0.08	0.08	0.08	0.38

ステップ 1 でエネルギーを受信すると、ノード B、C、および D がステップ 2 でアクティブ化し、ターゲットの近隣にエネルギーを拡散します。ノード B にはターゲット ノードがないため、エネルギーは拡散しません。それに対しノード C にはターゲットがありエネルギーの 80% が G に拡散します。そしてノード D は 80% のエネルギーをノード E および F に拡散、D との関係が強いため、E のエネルギーの 3 倍受信します。

ステップ 3 で、ノード E、F、G がアクティブ化します。ノード E および G にはターゲット ノードがないため、いずれもエネルギーを拡散しません。ただし、ノード F は 80% のエネルギーを 1 つのターゲット ノード、G に拡散します。このノードの総エネルギーは 0.38 で、ノード C および F から拡散したエネルギーで構成されています。

この時点で、ノード G がアクティブ化しますが、拡散のターゲット ノードがないため、プロセスが終了します。ノード A から拡散するエネルギーはネットワーク全体で拡散し、ノード G が最も多いエネルギーを受信しま

す。ネットワークに導入されたエネルギーがチェーンを示す場合、ノード G はノード A のチェーンの影響が最も大きくなります。

拡散分析ノード

[ソース] パレットで使用できる拡散分析ノードは、ソーシャル ネットワーク全体の指定された個人のセットからの効果を拡大し、ネットワークの関係を使用して、効果に最も影響を受けている個人を特定します。例えば効果が解約の場合、ノードは解約するネットワーク内の他の特定の個人によって解約する可能性が最も高い個人を特定します。既存のモデルとデータをノード出力によって増補し、これらのモデルによって生成された予測を改善することができます。たとえば、個人の解約の傾向を更新して拡散の影響を含めることができます。

要件: ノードには分析するデータを含む 2 つの固定幅のテキスト ファイルが必要です。1 つ目のファイルは、3 つのフィールドを使用してソーシャル ネットワークを定義します。1 つ目のフィールドは方向が指定された各関係のソースを定義、2 つ目のフィールドは方向が指定された各関係の宛先を定義、3 つ目のフィールドは各関係の任意の強度を指定します。ネットワーク内のすべての関係には方向を指定する必要があります。2 つ目のファイルには、効果が開始する識別子のリストが含まれています。

拡散分析のデータの指定

[拡散分析] ソース ノード ウィンドウの [データ] タブを使用して、ネットワーク ノードの関係を指定する入力ファイルを指定できます。

ファイル: 通話詳細レコードを含むファイルまたはフォルダの名前を指定します。名前を入力するか、[...] ボタンをクリックしてファイル システムの名前を選択します。名前を選択すると、下部にファイル パスが表示されます。また、ファイルの内容が、区切り文字付きで表示されます。フォルダを指定する場合、フォルダに含まれるすべてのファイルの呼び出しの詳細レコードが連結され、分析されます。フォルダのすべてのファイルの構造は同じでなければなりません。

ファイルからフィールド名を取得: デフォルトで選択されているこのオプションは、データ ファイル中の最初の行を列の名前として取り扱います。1 行目が見出しでない場合、このオプションをオフにして各フィールドに自動的に Field1 や Field2 といった汎用名を割り当てます。

ネットワーク定義設定

ネットワーク定義設定はフィールドの役割を定義します。

フィールド: 矢印ボタンを使用して、このリストから項目をさまざまな役割フィールドに手動で割り当てます。アイコンは、各役割フィールドの有効な測定レベルを示します。[すべて] ボタンをクリックしてリスト内のすべてのフィールドを選択するか、各測定レベルのボタンをクリックして、その測定レベルのすべてのフィールドを選択します。

ソース: 方向付き関係の起点としてフィールドを 1 つ選択します。

ターゲット: 方向付き関係のターゲットとしてフィールドを 1 つ選択します。

重み: 関係の重みを示すフィールドを 1 つ選択します。通話の詳細に関するレコードの場合、重みは通話時間または通話の頻度など、関係の強度を反映するさまざまな通話の特性に対応しています。

拡散分析のビルド オプションの設定

[拡散分析] ソース ノード ウィンドウの [ビルド オプション] タブを使用して、拡散プロセスのオプションを定義できます。

拡散シード リスト

拡散シード リストは、エネルギーが拡散し始めるネットワーク ノードを特定します。たとえば、解約分析に置いて、このリストはすでに解約したネットワーク ノードを特定します。また、このリストには、これらのノードが実際に解約するネットワークの効果を探るために、解約のリスクが最も高いノードを含めることができます。

ファイル: ノードの識別子を含むテキスト ファイルの名前を指定します。各識別子はファイル内の各行に対応します。ファイル名を入力するか、[...] ボタンをクリックしてファイルを選択します。

拡散分析設定

拡散分析設定により、拡散の強度およびプロセス停止の時期を指定します。

拡散係数: 拡散プロセスの各ステップでノードが拡散するエネルギーの割合を定義します。値が小さいほど、エネルギーを拡散しているノードに最も近いノードにエネルギーを制限し、値が大きいほど、よりは慣れたノードに影響を与えることができます。

最大反復回数: 拡散の反復回数の上限を指定します。この上限に達した場合、自動的に拡散プロセスが終了します。

精度しきい値: 拡散プロセスの続行を保証する、ノードのエネルギーの変化の最小値を指定します。すべてのノードのエネルギーの変化がこの値より小さい場合、拡散プロセスが終了します。

要約統計量の計算および表示: このオプションを選択すると、ノードは主要業績評価指標出力を生成するほか、分析の要約統計量を計算して表示します。ネットワークが極めて大きい場合、これらの統計量の計算は、ノードのパフォーマンスに悪影響を与える場合があります。

拡散分析統計の表示

[拡散分析] ソース ノード ウィンドウの [分析] タブには、拡散の結果の概要が表示されます。“拡散分析の要約統計” 表には、利用できる要約統計を示します。

テーブル 3-2
拡散分析の要約統計

統計	説明
ネットワーク内のノード数	ネットワーク内のノード数
ネットワーク内のリンク数	ネットワーク内のリンク数
ネットワーク内の拡散シード数	拡散プロセスのシードとして使用されるノード数
平均の影響	個人と関連する拡散エネルギーの平均量
平均 InDegree	個人が関係のターゲットである関係の平均数 詳細は、1 章 p.5 ノードの度数 を参照してください。
平均 OutDegree	個人が関係のソースである関係の平均数 詳細は、1 章 p.5 ノードの度数 を参照してください。

要約統計表から特定の統計を選択して、値の分布、標準偏差、統計の歪度を表示します。

拡散分析の出力

拡散分析ノードは、ネットワーク内の個人について説明するさまざまなフィールドを生成します。既存のモデルとデータを主要行政指標によって増補し、これらのモデルによって生成された予測を改善することができます。たとえば、個人の解約の傾向を更新して拡散の影響を含めることができます。

“拡散分析の主要業績指標” 表に、分析における個人の主要業績指標フィールドを示します。

テーブル 3-3
拡散分析の主要業績指標

フィールド	説明
DA_NodeNumber	個人の一意の識別子
DA_DiffusedEnergy	個人と関連する拡散エネルギーの量解約分析の場合、大きい値は小さい値に比べて解約する傾向が大きくなります。
DA_InDegree	個人が関係のターゲットである関係の数 詳細は、1章 p.5 ノードの度数を参照してください。
DA_OutDegree	個人が関係のソースである関係の数 詳細は、1章 p.5 ノードの度数を参照してください。

注意事項

この情報は、世界各国で提供される製品およびサービス向けに作成されています。

IBMはこのドキュメントで説明する製品、サービス、機能は他の国では提供していない場合があります。現在お住まいの地域で利用可能な製品、サービス、および、情報については、お近くの IBM の担当者にお問い合わせください。IBM 製品、プログラム、またはサービスに対する参照は、IBM 製品、プログラム、またはサービスのみが使用することができることを説明したり意味するものではありません。IBM の知的所有権を侵害しない機能的に同等の製品、プログラム、またはサービスを代わりに使用することができます。ただし、IBM 以外の製品、プログラム、またはサービスの動作を評価および確認するのはユーザーの責任によるものです。

IBMは、本ドキュメントに記載されている内容に関し、特許または特許出願中の可能性があります。本ドキュメントの提供によって、これらの特許に関するいかなる権利も使用者に付与するものではありません。ライセンスのお問い合わせは、書面にて、下記住所に送ることができます。

IBM Director of Licensing, IBM Corporation, North Castle Drive,
Armonk, NY 10504-1785, U. S. A.

以下の条項はは、イギリスまたはこのような条項が法律に反する他の国では適用されません。 International Business Machines は、明示的または黙示的に関わらず、第三者の権利の侵害しない、商品性または特定の目的に対する適合性の暗黙の保証を含むがこれに限定されない、いかなる保証なく、本出版物を「そのまま」提供します一部の州では、特定の取引の明示的または暗示的な保証の免責を許可していないため、この文が適用されない場合があります。

この情報には、技術的に不適切な記述や誤植を含む場合があります。情報については変更が定期的に行われます。これらの変更は本書の新版に追加されます。IBM は、本書に記載されている製品およびプログラムについて、事前の告知なくいつでも改善および変更を行う場合があります。

IBM 以外の Web サイトに対するこの情報内のすべての参照は、便宜上提供されているものであり、決してそれらの Web サイトを推奨するものではありません。これらの Web サイトの資料はこの IBM 製品の資料に含まれるものではなく、これらの Web サイトの使用はお客様の責任によるものとします。

IBM はお客様に対する一切の義務を負うことなく、自ら適切と考える方法で、情報を使用または配布することができるものとします。

本プログラムのライセンス取得者が (i) 別途作成されたプログラムと他のプログラム (本プログラムを含む) との間の情報交換および (ii) 交換された情報の相互利用を目的とした本プログラムに関する情報の所有を希望する場合、下記住所にお問い合わせください。

IBM Software Group, Attention:Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

上記のような情報は、該当する条項および条件に従い、有料で利用できるものとしします。

本ドキュメントに記載されている許可されたプログラムおよびそのプログラムに使用できるすべてのライセンス認証された資料は、IBM Customer Agreement、IBM International Program License Agreement、および当社とかわした同等の契約の条件に基づき、IBM によって提供されます。

この情報には、日常の業務処理で用いられるデータや報告書の例が含まれています。できる限り詳細に説明するため、例には、個人、企業、ブランド、製品などの名前が使用されています。これらの名称はすべて架空のものであり、実際の企業で使用される名称および住所とは一切関係ありません。

商標

IBM、IBM ロゴ、および [ibm.com](http://www.ibm.com)、SPSS は、世界の多くの国で登録された IBM Corporation の商標です。IBM の商標の現在のリストは、<http://www.ibm.com/legal/copytrade.shtml> を参照してください。

Linux は、米国およびその他の国における Linus Torvalds の登録商標です。

Microsoft、Windows、Windows NT、および Windows のロゴは、米国およびその他の国における Microsoft 社の商標です。

UNIX は、米国およびその他の国における The Open Group の登録商標です。その他の製品名およびサービス名等は、IBM または他の会社の商標です。

用語集

Indegree. ノードがターゲットとなる関係数の合計として定義される、方向付き関係で構成されたネットワーク内のノードの地位の指標。

Outdegree. ノードがソースとなる関係数の合計として定義される、方向付き関係で構成されたネットワーク内のノードの中心性の指標。

カバレッジしきい値. グループの識別に使用される最も強力な関係の割合。

コアグループ. 弱い関係をネットワークから除外し、グループ サイズの制限を適用した後に残ったノードのグループ。

ソシオグラム. 点を繋ぐ線がお互いに関係を持つ個人を示すソーシャルネットワークの視覚的表現。

二分関係. 繋がりの有無を示す、2 つの値の一方のみを使用する 2 ノード間の繋がり。

密度. 観測された関係数 ÷ 可能な関係の数として定義される、ネットワークの結合の指標。

度数. ノードが関わる関係数の合計として定義される、ネットワークのノードの活動の指標。

拡散エネルギー. ネットワーク全体で拡散するネットワークの特定のノードを起点とする効果。最初のノードからの距離が大きくなるとサイズが小さくなります。

拡散スコア. 個人がグループの他のメンバーにどれだけ接触しているかを示す指標。拡散スコアは、個人の重要度をその人物がソースとなる関係の数と関連付けます。

拡散リーダー. 拡散スコアが最も高いグループの個人。

拡散係数. 拡散プロセス ステップでノードが近隣に送信するエネルギーの割合。

方向なし関係. 2 つのノード間の関連性。

方向付き関係. 一方のノード（ソース）から始まり、もう一方（ターゲット）で終わる 2 つのノード間の接続。

権限スコア. グループのほかのメンバーが個人にどれだけ接触しているかを示す指標。権限スコアは、個人の重要度をその人物がターゲットとなる関係の数と関連付けます。

権限リーダー. 権限スコアが最も高いグループの個人。

精度しきい値. 拡散の続行を保証するノードのエネルギーの最も小さい変化に対応する拡散プロセスの停止基準。

評価付き関係. 関係の強度を示す重みが割り当てられた 2 つのノード間の繋がり。

参考文献

Cover, T. M., および J. A. Thomas. 2006. Elements of Information Theory, 2nd edition. New York: John Wiley and Sons, Inc.

Dasgupta, K., R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjea, A. A. Nanavati, および A. Joshi. 2008. Social ties and their relevance to churn in mobile telecom networks. In: Proceedings of the 11th international conference on extending database technology, New York, NY: ADM, 668-677.

Moreno, J. L. 1934. Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama. Washington, D.C.: Nervous and Mental Disease Publishing Co..

Richter, Y., E. Yom-Tov, および N. Slonim. 2010. Predicting customer churn in mobile networks through analysis of social groups. In: Proceedings of the 2010 SIAM international conference on data mining, Columbus, OH: SDM 2010, 732-741.

Wasserman, S., および K. Faust. 1994. Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

索引

- 主要業績指標, 18, 24
- 二分関係, 2
- 拡散係数, 20, 23
- 拡散分析, 7, 21
 - 拡散係数, 20, 23
 - 要件, 22
 - オプション, 22
 - 精度しきい値, 20, 24
 - シード, 23
- 商標, 27
- 密度, 5, 17-18
- 尺度, 7
- 度数, 5
- 関係
 - 価値, 2
 - 方向, 2
- 平均 indegree, 17
- 平均 outdegree, 17, 24
- [分析] タブ, 7
 - 拡散分析, 24
 - グループ分析, 17
- [注釈] タブ, 7

- GAG_GroupNumber, 18
- indegree, 6, 17, 19, 24-25
- 平均indegree, 24
- outdegree, 6, 17, 19, 24-25

- 拡散エネルギー, 24-25

- カバレッジしきい値, 13, 16
- カーネル密度, 17-18

- 方向付き関係, 2
- 評価付き関係, 2

- グループ分析, 7
 - 要件, 15
 - カバレッジしきい値, 13, 16
 - グループ サイズ, 14, 16
- 平均グループ密度, 17
- グループ サイズ, 17-18
 - 上限, 14, 16
- 平均グループ サイズ, 17

- コア グループ, 14
- コア グループ密度の平均, 17
- コア ノード, 17-18
- コアメンバーの平均割合, 17

- 精度しきい値, 20, 24
- シード
 - 拡散分析, 23
- 拡散シード リスト, 23

- スクリプト
 - プロパティ, 10
- 拡散スコア, 14, 18-19
- 権限スコア, 14, 18-19
- ストリームの実行
 - IBM SPSS Modeler Server Social Network Analysis の使用, 7

- ソース フィールド, 9, 16, 23

- ターゲット フィールド, 9, 16, 23

- [データ型] タブ, 7
- [データ] タブ, 7
 - 拡散分析, 22
 - グループ分析, 15
- データの分析, 8

- 方向なし関係, 2

- 法律に関する注意事項, 26

- 平均の影響, 24
- 結果のクリア, 8
- 出力のフィルタリング, 7
- 出力のプレビュー, 8

- [ビルド オプション] タブ, 7
 - 拡散分析, 23
 - グループ分析, 16

- [フィルタ] タブ, 7
- プロパティ
 - スクリプト, 10

- 重みフィールド, 9, 16, 23

- 拡散リーダー, 14, 19
- 権限リーダー, 14, 19

- 通話詳細レコード, 9