

IBM SPSS Modeler Social Network
Analysis 15 – Benutzerhandbuch



Hinweis: Lesen Sie zunächst die allgemeinen Informationen unter Hinweise auf S. 24, bevor Sie dieses Informationsmaterial sowie das zugehörige Produkt verwenden.

Diese Ausgabe bezieht sich auf IBM SPSS Modeler Social Network Analysis 15 und alle nachfolgenden Versionen sowie Anpassungen, sofern dies in neuen Ausgaben nicht anders angegeben ist.

Screenshots von Adobe-Produkten werden mit Genehmigung von Adobe Systems Incorporated abgedruckt.

Screenshots von Microsoft-Produkten werden mit Genehmigung der Microsoft Corporation abgedruckt.

Lizenziertes Material - Eigentum von IBM

© **Copyright IBM Corporation 2012.**

Eingeschränkte Rechte für Benutzer der US-Regierung: Verwendung, Vervielfältigung und Veröffentlichung eingeschränkt durch GSA ADP Schedule Contract mit der IBM Corp.

IBM® SPSS® Modeler Social Network Analysis verarbeiten Informationen zu Beziehungen zwischen Personen in Felder, die die Rolle einer Einzelperson in einem sozialen Netzwerk beschreiben, wodurch soziale Informationen in Vorhersagemodelle aufgenommen werden können. In diesem Handbuch wird die Verwendung der IBM SPSS Modeler Social Network Analysis-Knoten in der IBM® SPSS® Modeler-Umgebung beschrieben, die Ihnen die Aufnahme der Knoten in Ihre Streams. Wenn die Knotenausgabe mit Feldern kombiniert wird, die für Maßzahlen zu Einzelpersonen stehen, ergibt sich ein vollständigeres Profil der Ergebnisse für die Einzelpersonen.

Informationen zu IBM Business Analytics

Die Software IBM Business Analytics liefert umfassende, einheitliche und korrekte Informationen, mit denen Entscheidungsträger die Unternehmensleistung verbessern können. Ein umfassendes Portfolio aus [Business Intelligence](#), [Vorhersageanalyse](#), [Finanz- und Strategiemangement](#) sowie [Analyseanwendungen](#) bietet Ihnen sofort klare und umsetzbare Einblicke in die aktuelle Leistung und gibt Ihnen die Möglichkeit, zukünftige Ergebnisse vorherzusagen. Durch umfassende Branchenlösungen, bewährte Vorgehensweisen und professionellen Service können Unternehmen jeder Größe die Produktivität maximieren, Entscheidungen automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt IBM SPSS Predictive Analytics-Software Unternehmen dabei, zukünftige Ereignisse vorherzusagen und proaktiv Maßnahmen zu ergreifen, um bessere Geschäftsergebnisse zu erzielen. Kunden aus Wirtschaft, öffentlichem Dienst und dem Bildungsbereich weltweit nutzen IBM SPSS-Technologie als Wettbewerbsvorteil für Kundengewinnung, Kundenbindung und Erhöhung der Kundenumsätze bei gleichzeitiger Eindämmung der Betrugsmöglichkeiten und Minderung von Risiken. Durch die Einbindung von IBM SPSS-Software in ihre täglichen Operationen wandeln sich Organisationen zu “Predictive Enterprises” – die Entscheidungen auf Geschäftsziele ausrichten und automatisieren und einen messbaren Wettbewerbsvorteil erzielen können. Wenn Sie weitere Informationen wünschen oder Kontakt zu einem Mitarbeiter aufnehmen möchten, besuchen Sie die Seite <http://www.ibm.com/spss>.

Technischer Support

Kunden mit Wartungsvertrag können den technischen Support in Anspruch nehmen. Kunden können sich an den technischen Support wenden, wenn sie Hilfe bei der Arbeit mit IBM Corp.-Produkten oder bei der Installation in einer der unterstützten Hardware-Umgebungen benötigen. Zur Kontaktaufnahme mit dem technischen Support besuchen Sie die Website von IBM Corp. unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihre Organisation und Ihre Supportvereinbarung angeben.

1 IBM SPSS Modeler Social Network Analysis 1

Informationen zu IBM SPSS Modeler Social Network Analysis	1
Analyse sozialer Netzwerke	2
Anzeigen von Netzwerken	3
Beschreiben von Netzwerken	4
IBM SPSS Modeler Social Network Analysis-Knoten	6
Knotenregisterkarten	7
Anzeigen einer Ausgabevorschau	7
Analysieren von Daten	8
Anwendungen	8
Datenstruktur	9
Skript-Eigenschaften	9

2 Gruppenanalyse 11

Gruppenanalyse – Übersicht	11
Ermitteln der sozialen Ähnlichkeit	11
Partitionierung in Gruppen	12
Beschreiben von Gruppen und Gruppenmitgliedern	13
Gruppenanalyseknoten	14
Angabe von Daten für die Gruppenanalyse	14
Festlegen von Erstellungsoptionen für die Gruppenanalyse	15
Anzeige von Statistiken zur Gruppenanalyse	16
Ausgabe für die Gruppenanalyse	16

3 Diffusionsanalyse 19

Diffusionsanalyse – Übersicht	19
Beispiel für Diffusionsanalyse	20
Diffusionsanalyseknoten	20
Angabe von Daten für die Diffusionsanalyse	21
Festlegen von Erstellungsoptionen für die Diffusionsanalyse	22
Anzeige von Statistiken zur Diffusionsanalyse	22
Ausgabe für die Diffusionsanalyse	23

Anhang

<i>A Hinweise</i>	<i>24</i>
<i>Glossar</i>	<i>26</i>
<i>Bibliografie</i>	<i>27</i>
<i>Index</i>	<i>28</i>

IBM SPSS Modeler Social Network Analysis

Informationen zu IBM SPSS Modeler Social Network Analysis

Zahlreiche Ansätze zur Modellierung von Verhalten konzentrieren sich die jeweilige Einzelperson. Dabei wird eine Reihe von Daten über Einzelpersonen verwendet, um ein Modell zu generieren, das Schlüsselindikatoren des Verhaltens verwendet, um dieses vorherzusagen. Wenn bei einer Einzelperson Werte für die Schlüsselindikatoren vorkommen, die dem Auftreten des Verhaltens zugeordnet sind, kann die Aufmerksamkeit gezielt auf diese Person gerichtet werden, um das betreffende Verhalten zu verhindern.

Betrachten Sie beispielsweise Ansätze zur Modellierung der Abwanderung, also wenn ein Kunde seine Geschäftsbeziehung mit einem Unternehmen beendet. Die Kosten, Kunden zu halten, sind bedeutend niedriger als die Kosten, sie zu ersetzen. Dies macht die Fähigkeit, abwanderungsgefährdete Kunden zu erkennen, so wichtig. Analysten verwenden häufig eine Reihe von Leistungskennzahlen (Key Performance Indicators, KIP), um Kunden zu beschreiben. Dazu gehören auch demografische Informationen und das jüngste Telefonieverhalten der einzelnen Kunden. Bei Vorhersagemodellen, die auf diesen Feldern beruhen, werden Änderungen im Telefonieverhalten der Kunden, die mit dem Telefonieverhalten von Kunden übereinstimmen, die früher abgewandert sind, verwendet, um Personen zu identifizieren, bei denen ein erhöhtes Abwanderungsrisiko besteht. Kunden, bei denen ein hohes Risiko ermittelt wurde, erhalten zusätzlichen Kundendienst oder Service-Optionen, um zu versuchen, sie zu halten.

Bei diesen Methoden werden soziale Informationen übersehen, die das Verhalten von Kunden erheblich beeinflussen können. Informationen über ein Unternehmen und darüber, was andere Personen tun, breiten sich über die Beziehungen aus und beeinflussen die Personen. Infolgedessen können Personen durch ihre Beziehungen zu anderen Personen die Entscheidungen und Handlungen der anderen beeinflussen. Analysen, die nur Maßzahlen für Einzelpersonen enthalten, vernachlässigen wichtige Faktoren, die Aussagekraft für die Prognose besitzen.

IBM® SPSS® Modeler Social Network Analysis begegnet diesem Problem, indem Beziehungsinformationen verarbeitet und in zusätzliche Felder eingetragen werden, die in Modelle aufgenommen werden können. Diese abgeleiteten Leistungskennzahlen messen soziale Merkmale für Einzelpersonen. Die Verbindung dieser sozialen Eigenschaften mit Maßzahlen für Einzelpersonen bietet einen besseren Überblick über Einzelpersonen und kann folglich die Vorhersagegenauigkeit Ihrer Modelle verbessern.

IBM SPSS Modeler Social Network Analysis besteht aus zwei Hauptkomponenten:

- Zur IBM® SPSS® Modeler-Umgebung hinzugefügte IBM SPSS Modeler Social Network Analysis-Knoten, die die Einbeziehung von Verfahren zur sozialen Analyse in Streams ermöglichen.
- IBM® SPSS® Modeler Server Social Network Analysis, der IBM® SPSS® Modeler Server um die Verarbeitung der Knotenspezifikationen ergänzt. IBM SPSS Modeler Server Social Network Analysis verarbeitet effizient enorme Mengen von Netzwerkdaten, die Millionen von Einzelpersonen und Beziehungen umfassen können, zu einer relativ kleinen Anzahl von Feldern zur weiteren Analyse.

So ermittelt IBM SPSS Modeler Social Network Analysis beispielsweise die Personen in einem Netzwerk, die am meisten durch die Abwanderung bestimmter Personen beeinflusst werden. Außerdem können Sie Gruppen von Personen in einem Netzwerk ermitteln, die einem erhöhten Abwanderungsrisiko unterliegen. Indem Sie Leistungskennzahlen für diese Effekte in Ihre Modelle aufnehmen, können sie deren Gesamtleistung verbessern.

Hinweis: Die Verarbeitung von IBM SPSS Modeler Social Network Analysis-Knoten muss durch IBM SPSS Modeler Server Social Network Analysis erfolgen. Wenn Sie einen Stream ausführen, der IBM SPSS Modeler Social Network Analysis-Knoten in SPSS Modeler enthält, müssen Sie eine Verbindung zu einer SPSS Modeler Server-Instanz herstellen, die ebenfalls IBM SPSS Modeler Server Social Network Analysis umfasst.

Analyse sozialer Netzwerke

Soziale Netzwerke bestehen aus einer Menge von Einzelpersonen und den Beziehungen zwischen diesen Personen. Bei der Analyse sozialer Netzwerke werden diese Beziehungen untersucht, um Einzelpersonen und Gruppen als Bestandteile einer sozialen Struktur zu beschreiben. Einzelpersonen interagieren miteinander, und diese Interaktionsmuster bieten Einblicke über die beteiligten Personen. Beziehungen ermöglichen den Informationsfluss in einem Netzwerk, sodass eine Person eine andere beeinflussen kann. Die Wichtigkeit der Beziehungsinformationen hebt die Analyse sozialer Netzwerke von anderen Ansätzen ab. Anstatt sich jeder Person einzeln zuzuwenden, ist die Untersuchungseinheit eine Dyade, bestehend aus zwei Einzelpersonen und ihren Beziehungen.

Beziehungen in einem Netzwerk können in gerichtete und ungerichtete Beziehungen unterteilt werden. Bei einer **gerichteten Beziehung** wird eine Person als Initiator bzw. Quelle der Beziehung gekennzeichnet und die andere als Empfänger bzw. Ziel. Ein Telefonanruf ist beispielsweise eine gerichtete Beziehung, bei der eine Person eine andere anruft. Im Gegensatz dazu können bei **ungerichteten Beziehungen** die Rollen "Quelle und Ziel" nicht definiert werden. In diesem Fall beteiligen sich beide Parteien gleichermaßen an der Beziehung. Ein Gespräch miteinander ist ein Beispiel für eine ungerichtete Beziehung.

Eine weitere Eigenschaft zur Unterscheidung von Beziehungen besteht darin, ob die Beziehung dichotom oder bewertet ist. Bei **dichotomen Beziehungen** ist die einzige verfügbare Information, ob die Beziehung zwischen zwei Personen besteht oder nicht. Für jede Dyade im Netzwerk ist die Beziehung entweder vorhanden oder abwesend. **Bewertete Beziehungen** dagegen beinhalten ein

Gewicht, das die Stärke der Beziehung anzeigt. Durch die Gewichte können die Beziehungen miteinander verglichen werden.

In der Tabelle “**Beziehungstypen**” wird ein Beispiel für die Kreuzklassifizierung von Beziehungen anhand von Richtung und Skala aufgeführt. Bei den gerichteten Beziehungen ist *Hans* die Quelle der Beziehung und *Marie* das Ziel. Bei den ungerichteten Beziehungen gibt es keinen Hinweis darauf, wer die Beziehung initiiert hat. Bei den bewerteten Beziehungen wird die Länge der Konversation als Beziehungsgewicht verwendet, während die dichotomen Beziehungen entweder auftraten oder nicht.

Tabelle 1-1
Beziehungstypen

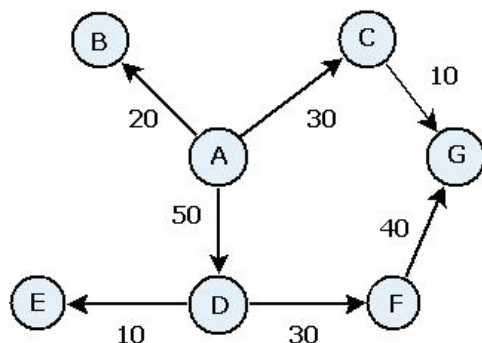
Richtung	Skala	Beispiel
Ungerichtet	Dichotom	Hans und Marie sprachen miteinander.
Ungerichtet	Bewertet	Hans und Marie sprachen 20 Minuten lang miteinander.
Gerichtet	Dichotom	Hans rief Marie an.
Gerichtet	Bewertet	Hans rief Marie für ein 20-minütiges Gespräch an.

Weitere Informationen zur Analyse sozialer Netzwerke finden Sie in den ausführlichen Büchern zu diesem Thema (Wasserman als auch Faust, 1994).

Anzeigen von Netzwerken

Soziale Netzwerke werden üblicherweise mithilfe von **Soziogrammen**(Moreno, 1934) dargestellt. Bei dieser visuellen Darstellung entsprechen die einzelnen Personen Punkte bzw. Knoten in einem Raum. Linien bzw. Kanten, die die Punkte verbinden, stellen Beziehungen zwischen den Einzelpersonen dar. Wenn die Beziehungen gerichtet sind, enthalten die Kanten einen Pfeil zur Angabe der Richtung. Wenn die Beziehungen Gewichte aufweisen, werden die entsprechenden Werte durch Beschriftungen für die Kanten angegeben. Die folgende Grafik zeigt ein Netzwerk für sieben Einzelpersonen.

Abbildung 1-1
Beispiel für soziales Netzwerk



Angenommen, das Netzwerk stellt die von einzelnen Personen getätigten Telefonanrufe dar, wobei die Beziehungsgewichte die Länge der Anrufe angeben. In diesem Fall rief Person A drei Personen an, wobei sie den größten Teil der Zeit mit Person D sprach.

Dieses Netzwerk ist deutlich kleiner als die in der Praxis vorkommenden Netzwerke. Die durch einfache Soziogrammen dargestellten Konzepte können jedoch als Verallgemeinerung von Netzwerken jeder Größe und Komplexität dienen.

Beschreiben von Netzwerken

Informationen zu Netzwerken, Gruppen und Einzelpersonen müssen in beschreibende Merkmale extrahiert werden, die Kreuzvergleiche und die Aufnahme in Vorhersagemodelle gestatten. Netzwerke müssen zu einer endlichen Menge von analysierbaren Leistungskennzahlen zusammenschmolzen werden. Beispielsweise kann es erforderlich sein, Netzwerke oder Gruppen von Knoten in einem Netzwerk miteinander zu vergleichen. Alternativ kann es sinnvoll sein, Einzelpersonen im Netzwerk miteinander zu vergleichen, um die wichtigsten Personen zu ermitteln.

Zwei häufig zur Beschreibung sozialer Netzwerke verwendete Maßzahlen sind **Dichte** und **Grad**. Beide Statistiken geben die Konnektivität an, die erstere bezieht sich jedoch auf das gesamte Netzwerk bzw. auf Netzwerkuntergruppen, wohingegen die letztere die Einzelpersonen im Netzwerk beschreibt.

Netzwerkdichte

Für jede gegebene Menge an Knoten in einem Netzwerk ist stets eine endliche Anzahl an Beziehungen möglich. Die einzelnen Knoten können als Quelle oder Ziel einer Beziehung mit jedem anderen Knoten dienen. Betrachten wir ein Netzwerk, das aus drei Knoten, A, B und C besteht. In der folgenden Tabelle sind alle möglichen gerichteten Beziehungen zwischen den Knoten aufgeführt.

Tabelle 1-2
Mögliche gerichtete Beziehungen für drei Knoten

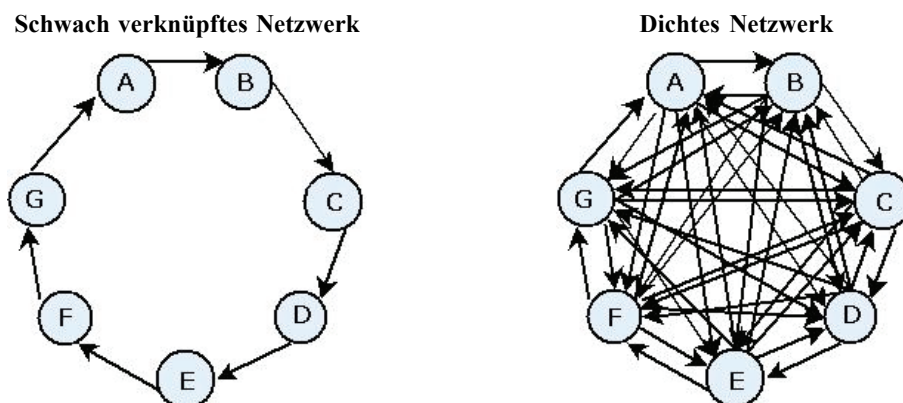
Quelle	Ziel
A	B
A	C
B	A
B	C
C	A
C	B

Jeder Knoten ist die Quelle einer Beziehung mit den beiden anderen Knoten. In der Praxis müssen jedoch nicht alle möglichen Beziehungen tatsächlich vorhanden sein. Einige Knoten weisen möglicherweise überhaupt keine direkte Beziehung mit anderen Knoten auf. Außerdem werden manche gerichteten Beziehungen möglicherweise nicht erwidert.

Die **Dichte**-Statistik sieht für den Anteil der möglichen Beziehungen im Netzwerk, die tatsächlich vorhanden sind. Der Wert reicht von 0 bis 1, wobei die Untergrenze einem Netzwerk ohne jegliche Beziehungen entspricht und die Obergrenze einem Netzwerk mit allen möglichen Beziehungen. Je näher der Wert an 1 liegt, desto dichter ist das Netzwerk und desto höher ist die Kohäsion der Knoten im Netzwerk.

Informationen können in dichten Netzwerken leichter fließen als in schwach verknüpften Netzwerken. Die Tabelle “Schwach verknüpfte und dichte Netzwerke” zeigt zwei Netzwerke an, die jeweils aus sieben Knoten bestehen. Das schwach verknüpfte Netzwerk enthält nur sieben der möglichen 42 Beziehungen zwischen den Knoten. Dies ergibt eine Dichte von 0,17. Das dichte Netzwerk dagegen enthält alle möglichen Beziehungen und weist eine Dichte von 1,0 auf.

Tabelle 1-3
Schwach verknüpfte und dichte Netzwerke



Im schwach verknüpften Netzwerk müssen Informationen fünf andere Knoten passieren, um von Knoten A zu Knoten G zu gelangen. Im Gegensatz dazu können die Informationen im dichten Netzwerk direkt von Knoten A zu Knoten G fließen.

Knotengrad

Die wichtigsten Personen in einem Netzwerk sind häufig diejenigen, die an der größten Anzahl an Beziehungen beteiligt sind. Diese Einzelpersonen erlangen Informationen aus einer Vielzahl von Quellen und verbreiten diese Informationen an eine große Anzahl anderer Personen weiter. Im Gegensatz dazu sind Personen, die nur an wenigen Beziehungen beteiligt sind, nicht in der Lage, eine große Anzahl anderer Personen im Netzwerk zu beeinflussen.

Der **Grad** eines Knotens, der als Gesamtzahl der Beziehungen definiert ist, an denen der Knoten beteiligt ist, ermöglicht Vergleiche zwischen den Netzwerkteilnehmern. Einzelpersonen mit höheren Gradwerten sind aktiver als Einzelpersonen mit niedrigeren Werten. Beim Gradwert (engl. “degree”) wird die Richtung der Beziehungen ignoriert. Es wird lediglich ein Gesamtmaß der Aktivität für den Knoten angegeben.

Bei gerichteten Beziehungen können Sie bei der Zählung der Beziehungen berücksichtigen, ob es sich bei einem Knoten um Quelle oder Ziel handelt. Der Wert **Indegree** für einen Knoten ist die Anzahl der Beziehungen, bei denen der betreffende Knoten das Ziel ist. Entsprechend gibt **Outdegree** die Anzahl der Beziehungen an, bei denen der betreffende Knoten als Quelle fungiert. In der folgenden Tabelle werden die Grad-, Indegree- und Outdegree-Werte für die einzelnen Knoten in der Abbildung “[Beispiel für soziales Netzwerk](#)” angegeben.

Tabelle 1-4
Beispielwerte für Grad, Indegree und Outdegree

Knoten	Grad	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” gilt häufig als Maßzahl für das Ansehen bzw. Prestige. Bei höheren Indegree-Werten endet eine größere Anzahl an Beziehungen bei dem betreffenden Knoten. Anders ausgedrückt: Die betreffenden Personen werden von einer großen Anzahl von anderen Personen kontaktiert. Viele andere Knoten initiieren Beziehungen mit dem betreffenden Knoten. Outdegree dagegen gilt als Maßzahl der Zentralität. Bei höheren Werten hat eine höhere Anzahl an Beziehungen ihren Ursprung bei diesem Knoten. Die betreffenden Personen nehmen Kontakt mit einer großen Anzahl an anderen Personen auf.

Bei den Knoten im Beispielnetzwerk geben die Gradwerte an, dass die Knoten A und D die höchste und die Knoten B und E die niedrigste Aktivität aufweisen. Die Indegree-Werte zeigen, dass Knoten G das höchste Prestige aufweist. Gemäß den Outdegree-Werten ist Knoten A am zentralsten.

IBM SPSS Modeler Social Network Analysis-Knoten

Neben den vielen Standardknoten, die im Lieferumfang von IBM® SPSS® Modeler enthalten sind, können Sie auch mit IBM® SPSS® Modeler Social Network Analysis-Knoten arbeiten, um die Ergebnisse von Analysen sozialer Netzwerke in Ihre Streams mit aufzunehmen. In der Tabelle “IBM SPSS Modeler Social Network Analysis-Knoten” werden diese Knoten beschrieben, die in der Palette “Datenquellen” gespeichert sind.

Tabelle 1-5
IBM SPSS Modeler Social Network Analysis-Knoten

Knoten	Symbol	Beschreibung
Gruppenanalyse		Der Gruppenanalyseknoten importiert Daten aus Kommunikationsdatensätzen aus einer Textdatei mit fester Feldlänge, ermittelt Gruppen von Knoten innerhalb des Netzwerks, die von den Datensätzen definiert sind, und generiert Leistungskennzahlen für die Gruppen und Einzelpersonen im Netzwerk. Für weitere Informationen siehe Thema Gruppenanalyse – Übersicht in Kapitel 2 auf S. 11.
Diffusionsanalyse		Der Diffusionsanalyseknoten importiert Daten aus Kommunikationsdatensätzen aus einer Textdatei mit fester Feldlänge, verbreitet einen durch die Datensätze definierten Effekt im Netzwerk und generiert Leistungskennzahlen, die die Ergebnisse des Effekts auf einzelne Knoten zusammenfassen. Für weitere Informationen siehe Thema Diffusionsanalyse – Übersicht in Kapitel 3 auf S. 19.

Die Verarbeitung der IBM SPSS Modeler Social Network Analysis-Knoten muss durch IBM® SPSS® Modeler Server Social Network Analysis erfolgen. Wenn Sie einen Stream ausführen, der diese Knoten in SPSS Modeler enthält, müssen Sie eine Verbindung zu einer IBM® SPSS® Modeler Server-Instanz herstellen, die ebenfalls IBM SPSS Modeler Server Social Network Analysis umfasst.

Knotenregisterkarten

Die Knoten von IBM® SPSS® Modeler Social Network Analysis bieten folgende Registerkarten zur Definition und Vorschau der Analyse:

- **Registerkarte "Daten"**. Dient zur Identifikation der Datei mit den Informationen zum sozialen Netzwerk.
- **Registerkarte "Erstellungsoptionen"** Dient zur Definition der Einstellungen für die Analysen.
- **Registerkarte "Analyse"**. Dient zur Anzeige einer Übersicht der vorläufigen Ausgaben, die Hinweise dafür bietet, wie die Eingabeeinstellungen auf der Registerkarte "Daten" so geändert werden können, dass sie die optimalen Ergebnisse liefern.

Außerdem bieten die Knoten folgende Registerkarten, die bei alle Quellknoten in IBM® SPSS® Modeler zu finden sind:

- **Registerkarte "Filter"**. Dient zum Entfernen oder Umbenennen der von den Knoten erstellten Ausgabefelder. Die Registerkarte bietet dieselben Funktionen wie der Filterknoten.
- **Registerkarte "Typen"**. Dient zum Festlegen der Messniveaus für die von den Knoten erstellten Ausgabefelder. Die Registerkarte bietet dieselben Funktionen wie der Typknoten.
- **Registerkarte "Anmerkungen"**. Dient zum Umbenennen von Knoten, zur Bereitstellung benutzerdefinierter QuickInfos und zum Speichern ausführlicher Anmerkungen.

Weitere Informationen zu diesen allgemeinen Registerkarten finden Sie in der SPSS Modeler-Dokumentation.

Anzeigen einer Ausgabevorschau

Typischerweise wird die Ausgabe der Knoten als Eingabe für ein Vorhersagemodell verwendet. Wenn Sie die Ausgabe anzeigen möchten, könnten Sie einen Tabellenknoten zum Stream hinzufügen. Die Datenmenge ist jedoch extrem groß, was die Tabellenerstellung zeitaufwendig macht. Außerdem wird ihre Brauchbarkeit durch die Anzahl der Zeilen in der Tabelle begrenzt.

Als Alternative können Sie eine Vorschautabelle generieren, die eine Stichprobe der zu erstellenden Ausgabe anzeigt. Die Vorschau zeigt die generierten Felder für eine begrenzte Anzahl an Zeilen. Die Anzahl an Zeilen ist in den Stream-Eigenschaften definiert. Führen Sie zur Vorschau der Knotenausgabe folgende Schritte aus:

1. Öffnen Sie den Knoten.
2. Geben Sie die Dateneinstellungen auf der Registerkarte "Daten" an.
3. Definieren Sie die Analyseparameter auf der Registerkarte "Erstellungsoptionen".

4. Klicken Sie auf Vorschau.

Es wird ein Vorschauenfenster mit den Ergebnissen angezeigt. Außerdem wird durch die Vorschau der Ausgabe die Registerkarte "Analyse" des Knotens mit einem zusammenfassenden Überblick über die Ergebnisse ausgefüllt.

Analysieren von Daten

Die Analyseeinstellungen, die zu optimalen Ergebnissen führen, werden normalerweise in einem iterativen Verfahren ermittelt. Sie definieren die Einstellungen, führen die Analyse aus und überprüfen die Ergebnisse. Wenn die Ergebnisse nicht die gewünschte Aussagekraft aufweisen, ändern Sie die Einstellungen und führen die Analyse erneut durch.

Führen Sie zur Analyse der Eingabedaten für den Knoten folgende Schritte aus:

1. Öffnen Sie den Knoten.
2. Geben Sie die Dateneinstellungen auf der Registerkarte "Daten" an.
3. Definieren Sie die Analyseparameter auf der Registerkarte "Erstellungsoptionen".
4. Klicken Sie auf Daten analysieren.

Wenn in den Erstellungsoptionen angegeben ist, dass Übersichtsstatistiken angezeigt werden sollen, werden die Ergebnisse auf der Registerkarte "Analyse" angezeigt.

Wenn Sie eine Analyse wiederholen möchten, klicken sie auf Analyse löschen, um die aktuellen Ergebnisse zu löschen, bevor Sie auf Daten analysieren klicken.

Anwendungen

Hier einige Anwendungen, bei denen IBM® SPSS® Modeler Social Network Analysis besonders vorteilhaft sein kann:

- **Abwanderungsprognose.** Gruppenmerkmale können einen Einfluss auf die Abwanderungsquoten haben. Durch Konzentration auf Personen in Gruppen, bei denen ein erhöhtes Abwanderungsrisiko besteht, kann die Abwanderung möglicherweise vermieden werden. Außerdem können Sie Personen ermitteln, die ein hohes Abwanderungsrisiko aufgrund des Informationsflusses aufweisen, der von bereits abgewanderten Personen ausgeht.
- **Nutzen von Gruppenführern.** Gruppenführer üben einen starken Einfluss auf andere Gruppenmitglieder aus. Wenn die Abwanderung eines Gruppenführers vermieden werden kann, verringert sich vermutlich auch die Abwanderungsrate bei den Gruppenmitgliedern. Umgekehrt kann der Versuch, einen Gruppenführer von einem Mitbewerber abzuwerben, die Abwanderungsquote der Gruppenmitglieder erhöhen, die dem betreffenden Mitbewerber zuzurechnen sind und gleichzeitig die Abwanderungsquote der Ihrem Unternehmen zuzurechnenden Gruppenmitglieder verringern.
- **Marketing.** Gruppenführer können zur Einführung neuer Produkte oder Serviceangebote verwendet werden. Der Einfluss des Gruppenführers kann die Wahrscheinlichkeit erhöhen, dass andere Gruppenmitglieder das Angebot erwerben. Mithilfe der Diffusionsanalyse können

Sie die Personen ermitteln, die am meisten von den Gruppenführern beeinflusst werden und Ihre Marketingstrategien gezielt auf diese ausrichten.

Datenstruktur

Informationen zu den Personen in dem zu untersuchenden Netzwerk können über eine Vielzahl von Dateien, Datenbanken und Systeme in Ihrem gesamten Unternehmen verteilt sein. Um das Netzwerk mithilfe von IBM® SPSS® Modeler Social Network Analysis analysieren zu können, müssen Sie die relevanten Datensätze und Felder aus den Datenquellen extrahieren und sie für die Eingabe in die Knoten formatieren.

Die Analyseknotten benötigen Kommunikationsdatensätze, die in einer einzelnen Textdatei mit fester Breite gespeichert sind. Jede Zeile der Datei entspricht einer Beziehung. Dabei sind die Daten in folgenden Spalten organisiert:

- Die Kennung für die Person, die die Beziehung initiiert.
- Die Kennung für die Person, die das Ziel der Beziehung ist.
- Eine optionale Gewichtung für die Beziehung.

Alle Daten müssen numerisch sein, wobei als Kennungen für Einzelpersonen nur ganze Zahlen in Frage kommen. Optional können Feldnamen als erste Datenzeile der Datei mit aufgenommen werden. Diese Datenstruktur wird in der Tabelle [“Beispiele für Kommunikationsdatensätze”](#) verdeutlicht.

Tabelle 1-6
Beispiele für Kommunikationsdatensätze

Quelle	Ziel	Stärke
1000	5642	243
2190	8444	831
0299	9419	559

Die Gewichtswerte können jeder beliebigen Maßzahl entsprechen, die Sie verwenden möchten, um die Bedeutung der Beziehung in Bezug auf die anderen Beziehungen darzustellen. Bei Anrufrufen sind übliche Gewichte die Anrufdauer bzw. die Anruhfrequenz.

Wenn sich die Analyse auf eine Teilmenge des Anrufverlaufs beziehen soll, müssen Sie bei der Erstellung der Eingabedatei die betreffende Teilmenge verwenden. Sie können die Analyse beispielsweise auf die letzten Monate oder auf die aktuellsten Anrufe für eine Person eingrenzen, indem Sie nur diese Daten in die Eingabetextdatei aufnehmen.

Skript-Eigenschaften

Die Skripterstellung in IBM® SPSS® Modeler Social Network Analysis ist ein leistungsstarkes Tool, mit dem Prozesse in der Benutzeroberfläche automatisiert werden. Skripts können dieselben Arten von Aktionen durchführen, die Sie mit einer Maus oder einer Tastatur durchführen. So können Sie Aufgaben automatisieren, die bei einer manuellen Durchführung sehr viele Wiederholungen verlangen oder sehr viel Zeit beanspruchen. Weitere Informationen zur

Verwendung von Skripten finden Sie im Leitfaden *ScriptingAutomation.pdf*, der mit IBM® SPSS® Modeler bereitgestellt wird.

Eigenschaften für Diffusionsknoten

In der folgenden Tabelle sind die Skript-Eigenschaften für den Diffusionsknoten aufgeführt.

Eigenschaftsname	Datentyp	Eigenschaftsbeschreibung
input_data_file_name	<i>string</i>	
calling_field	<i>Feld</i>	
called_field	<i>Feld</i>	
frequency_weight_field	<i>Feld</i>	
read_field_names	<i>boolean</i>	
diffusion_list_file_name	<i>string</i>	
spreading_factor	<i>double</i>	Standardwert = 50. Minimalwert = 1 Maximalwert = 99.
max_number_iterations	<i>integer</i>	Standardwert = 100. Minimalwert = 1
accuracy_threshold	<i>double</i>	Standardwert = 0.01. Minimalwert = 0.001
calculate_statistics	<i>boolean</i>	

Eigenschaften für Gruppenanalyseknoten

In der folgenden Tabelle sind die Skript-Eigenschaften für den Gruppenanalyseknoten aufgeführt.

Eigenschaftsname	Datentyp	Eigenschaftsbeschreibung
input_data_file_name	<i>string</i>	
calling_field	<i>Feld</i>	
called_field	<i>Feld</i>	
frequency_weight_field	<i>Feld</i>	
read_field_names	<i>boolean</i>	
coverage_threshold	<i>double</i>	Standardwert = 10. Minimalwert = 1 Maximalwert = 99.
min_group_size	<i>integer</i>	Standardwert = 2. Minimalwert = 2
max_group_size	<i>integer</i>	Standardwert = 100. Minimalwert = 2
calculate_statistics	<i>boolean</i>	

Gruppenanalyse

Gruppenanalyse – Übersicht

Bei der Gruppenanalyse werden die Interaktionsmuster von Einzelpersonen in einem Netzwerk verwendet, um Gruppen ähnlicher Einzelpersonen zu ermitteln. Merkmale dieser Gruppen beeinflussen das Verhalten der einzelnen Gruppenmitglieder. Beispielsweise weisen kleine Gruppen mit vielen Beziehungen zwischen den Mitgliedern und starken Gruppenführern ein höheres Abwanderungsrisiko auf, selbst wenn kein Mitglied der Gruppe tatsächlich abgewandert ist. (Richter, Yom-Tov, als auch Slonim, 2010) Vorhersagemodelle, die sowohl Maßzahlen für Gruppen als auch Maßzahlen für Einzelpersonen beinhalten, weisen eine höhere Leistungsfähigkeit auf als Modelle, die nur die letzteren enthalten.

Die Gruppenanalyse besteht aus folgenden allgemeinen Schritten:

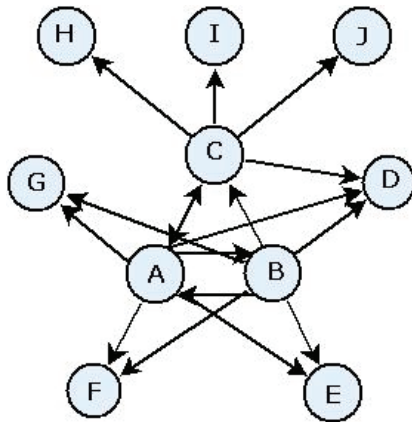
1. Ermitteln der Beziehungsstärken, die soziale Proximität widerspiegeln. Für weitere Informationen siehe Thema [Ermitteln der sozialen Ähnlichkeit](#) auf S. 11.
2. Partitionieren des Netzwerks in Gruppen auf der Grundlage der Beziehungsstärke (unter Berücksichtigung von Größenbeschränkungen). Für weitere Informationen siehe Thema [Partitionierung in Gruppen](#) auf S. 12.
3. Erstellen von Profilen für Gruppen und Einzelpersonen, einschließlich der Ermittlung von Gruppenführern. Für weitere Informationen siehe Thema [Beschreiben von Gruppen und Gruppenmitgliedern](#) auf S. 13.

Ermitteln der sozialen Ähnlichkeit

Die Mitglieder einer Gruppe sollten eine größere Ähnlichkeit zueinander aufweisen als zu Personen, die nicht zu der Gruppe gehören. Bei der Netzwerkanalyse hängt die Ähnlichkeit von zwei Knoten von ihren Beziehungen ab. Für jeden beliebigen Knoten in einem Netzwerk gibt es ein Set von Knoten, die die Ziele gerichteter Beziehungen mit dem betreffenden Knoten sind. Bei Telekommunikationsdaten entspricht dieses Set allen Personen, die von einer bestimmten Person kontaktiert wurden. Wenn zwei Personen dasselbe Set von Personen kontaktieren, gelten diese Personen als einander ähnlich. Je stärker sich die Sets der Beziehungsziele für zwei Personen überschneiden, desto ähnlicher sind sie.

Betrachten Sie das Netzwerk in der Abbildung [“Beispielnetzwerk mit zehn Knoten”](#).

Abbildung 2-1
Beispielnetzwerk mit zehn Knoten



In der Tabelle “Zielknoten” sehen Sie die Zielknoten für Beziehungen, die ihren Ursprung bei den Knoten A, B und C haben.

Tabelle 2-1
Zielknoten

Quellenknoten	Zielknoten
A	B, C, D, E, F, G
B	B, C, D, E, F, G
C	A, D, H, I, J

Die Knoten A und B haben fünf gemeinsame Zielknoten. Die Knoten A und C dagegen nur einen einzigen. Die Knoten A und B weisen also eine größere soziale Ähnlichkeit auf als die Knoten A und C.

Beziehungsgewichte, wie beispielsweise Anrufdauer oder -häufigkeit erfassen nicht die Ähnlichkeit der Knoten in einem Netzwerk. Um das Konzept der Ähnlichkeit mit aufzunehmen verwendet die Gruppenanalyse gegenseitige Informationen (Cover als auch Thomas, 2006) als Beziehungsgewicht. Diese Statistik gibt die Wahrscheinlichkeit an, dass zwei Knoten mit demselben Knoten verbunden sind. Wenn die Beziehungen im Netzwerk vordefinierte Gewichte aufweisen, wie beispielsweise die Anruhfrequenz bei Telekommunikationsdaten, schließen die gegenseitigen Informationen diese Gewichte entsprechend ein.

Partitionierung in Gruppen

Gruppen sollten aus Personen bestehen, die hohe Beziehungsgewichte mit anderen Gruppenmitgliedern aufweisen, wobei die Gewichte die Ähnlichkeit der Knoten in der Beziehung messen. Daher beginnt die Gruppenidentifikation durch Weglassen der schwächeren Beziehungen in einem Netzwerk. Der **Schwellenwert Abdeckung** steuert diesen Prozess, indem er den Anteil der stärksten Beziehungen angibt, der beibehalten werden soll. So führt beispielsweise ein Abdeckungsschwellenwert von 0,4 dazu, dass die stärksten 40 % der Beziehungen für die Gruppenidentifizierung verwendet werden, während die restlichen 60 % der Beziehungen weggelassen werden.

Die verbleibenden Beziehungen können einige sehr kleine oder sehr große Gruppen ergeben, die begrenzten Nutzen für Vorhersagen aufweisen. Um zu verhindern, dass solche Gruppen in die Analyse aufgenommen werden, können Sie Mindest- und Höchstgrößen für die Gruppen angeben. Gruppen, die kleiner sind als der Mindestwert, werden vollständig weggelassen. Gruppen, die größer sind als der Höchstwert, werden dagegen in kleinere Gruppen aufgeteilt, deren Größe innerhalb des zulässigen Bereichs liegt. Die nach dem Weglassen der schwachen Beziehungen und dem Durchsetzen der Größenbeschränkungen verbleibenden Gruppen werden als **Kerngruppen** bezeichnet.

Das Entfernen von Beziehungen aus dem ursprünglichen Netzwerk kann dazu führen, dass manche Personen zu keiner Kerngruppe gehören. Diese Personen weisen jedoch möglicherweise Beziehungen mit Mitgliedern einer Gruppe auf, die eine Aufnahme in die Gruppe rechtfertigen. Einzelpersonen werden zu Gruppen hinzugefügt, wenn es viele relativ starke Beziehungen mit den Kernmitgliedern der Gruppe gibt, vorausgesetzt die für Gruppen geltenden Größenbeschränkungen werden nicht verletzt. Die endgültigen Gruppen bestehen somit aus einer Menge von Kernmitgliedern sowie Mitgliedern, die aufgrund der Konnektivität mit dem Kern mit aufgenommen wurden.

Beschreiben von Gruppen und Gruppenmitgliedern

Neben Dichte, Indegree und Outdegree dienen noch weitere Statistikwerte zur Beschreibung von Gruppendynamiken. Insbesondere Autoritäts- und Disseminations-Scores bieten Maßzahlen für den sozialen Status der Personen innerhalb der Gruppen. Die Rolle der einzelnen Personen in einer Gruppe ist von entscheidender Bedeutung für den Versuch, das Verhalten der Gruppe und ihrer Mitglieder vorherzusagen.

Der **Autoritäts-Score** für einen Knoten misst die Tendenz anderer Knoten in der Gruppe, eine Verbindung mit ihm herzustellen. Wenn viele Personen eine bestimmte Person kontaktieren und möglicherweise um Informationen oder um die Meinung dieser Person bitten, kommt der betreffenden Person die Rolle einer Autorität zu. Die Autoritäts-Scores für die Knoten in der Gruppe, die den stationären Wahrscheinlichkeiten für eine Zufallsbewegung entsprechen, die erneut im Gruppennetzwerk beginnt, reichen von 0 bis 1. Je näher der Autoritäts-Score bei 1 liegt, desto größere Autorität hat der Knoten innerhalb der Gruppe. Der Knoten mit dem höchsten Autoritäts-Score innerhalb der Gruppe wird als **Autoritätsführer** der Gruppe bezeichnet. Die Division des größten Scores innerhalb einer Gruppe durch den kleinsten führt zu einer Maßzahl für die allgemeine Stärke des Autoritätsführers.

Im Gegensatz dazu misst der **Disseminations-Score** für einen Knoten die Tendenz des Knotens, eine Verbindung mit anderen Knoten in der Gruppe herzustellen. Wenn eine bestimmte Person mit vielen anderen Personen in der Gruppe Kontakt aufnimmt, kann diese Person einen erheblichen Einfluss auf die Meinungen der gesamten Gruppe haben. Die Disseminations-Scores für die Knoten in der Gruppe, die den stationären Wahrscheinlichkeiten für eine Zufallsbewegung entsprechen, die in umgekehrter Richtung erneut im Gruppennetzwerk beginnt, reichen von 0 bis 1. Je näher der Disseminations-Score bei 1 liegt, desto mehr stellt der Knoten Verbindungen mit den anderen Gruppenmitgliedern her. Der Knoten mit dem höchsten Disseminations-Score innerhalb der Gruppe wird als **Disseminationsführer** der Gruppe bezeichnet. Die Division

des größten Scores innerhalb einer Gruppe durch den kleinsten führt zu einer Maßzahl für die allgemeine Stärke des Disseminationsführers.

Gruppenanalyseknoten

Der Gruppenanalyseknoten, der über die Palette “Datenquellen” verfügbar ist, ermittelt Sets von Einzelpersonen in einem Netzwerk, die eine hohe soziale Ähnlichkeit untereinander aufweisen, und ermittelt den relativen Sozialstatus der Personen in den Gruppen. Gruppen, die relativ klein sind, zahlreiche Verbindungen zwischen den Mitgliedern aufweisen und einen starken Führer haben, sind aufgrund der Gruppendynamiken einem erhöhten Abwanderungsrisiko ausgesetzt. Sie können anhand der Ergebnisse Gruppen-Scores für die Abwanderung ermitteln, die in Kombination mit den Einzel-Scores für die Abwanderung die Fähigkeit zur Vorhersage der Abwanderung gegenüber Modellen erhöhen, die ausschließlich auf Einzelscores beruhen.

Angesichts der großen Datenmenge, aus denen das Netzwerk üblicherweise besteht, und der Natur der sozialen Interaktionen, die Gruppenelemente zugrunde liegen, ist die Gruppenanalyse ein zeitaufwendiger Vorgang. Üblicherweise werden die Ergebnisse der Gruppenanalyse in einer Datenbank oder einer Datei gespeichert, die als Eingangsdaten für ein Vorhersagemodell verwendet wird. Die Gruppenergebnisse werden in relativ unregelmäßigen Abständen aktualisiert, beispielsweise monatlich, wohingegen das Modell wesentlich häufiger aktualisiert werden kann.

Anforderungen. Für den Knoten ist eine Textdatei mit fester Breite erforderlich, die das soziale Netzwerk über drei Felder definiert. Ein Feld gibt die Quelle für jede gerichtete Beziehung an, ein Feld gibt das Ziel für jede gerichtete Beziehung an und das dritte Feld gibt eine optionale Stärke für jede Beziehung an. Alle Beziehungen im Netzwerk müssen gerichtet sein.

Angabe von Daten für die Gruppenanalyse

Auf der Registerkarte “Daten” des Fensters für den Gruppenanalyse-Quellknoten können Sie die Eingabedatei angeben, die die Beziehungen zwischen den Knoten im Netzwerk enthält.

Datei. Geben Sie den Namen der Datei oder des Ordners mit den Kommunikationsdatensätzen ein. Zur Auswahl eines Namens aus dem Dateisystem können Sie einen Dateinamen eingeben oder auf die Schaltfläche mit den Auslassungspunkten (...) klicken. Der Pfad wird angezeigt, sobald Sie einen Namen ausgewählt haben, und der entsprechende Inhalt wird mit Trennzeichen angezeigt. Wenn Sie einen Ordner angeben, werden die Kommunikationsdatensätze aller in dem Ordner befindlichen Dateien für die Analyse aneinandergereiht. Dabei müssen alle Dateien in dem Ordner die gleiche Struktur haben.

Feldnamen aus Datei lesen. Diese standardmäßig ausgewählte Option behandelt die erste Zeile der Datendatei als Namen für die Spalte. Wenn die erste Zeile keine Überschrift ist, deaktivieren Sie diese Option, um jedem Feld automatisch einen generischen Namen zuzuweisen, beispielsweise *Feld1* und *Feld2*.

Einstellungen für Netzwerkdefinition

Die Einstellungen für die Netzwerkdefinition legen die Rollen der Felder fest.

Felder. Verwenden Sie die Pfeilschaltflächen, um Elemente aus dieser Liste manuell den verschiedenen Rollenfeldern zuzuweisen. Die Symbole zeigen das gültige Messniveau für jedes Rollenfeld an. Klicken Sie auf die Schaltfläche **Alle**, damit alle Felder in der Liste ausgewählt werden, oder klicken Sie auf die Schaltfläche für ein individuelles Messniveau, um alle Felder mit diesem Messniveau auszuwählen.

Quelle. Wählen Sie ein Feld als Ursprung der gerichteten Beziehung aus.

Ziel. Wählen Sie ein Feld als Ziel der gerichteten Beziehung aus.

Gewicht. Optional können Sie ein Feld auswählen, das angibt, wie oft die Beziehung im Netzwerk vorkommt. Die Analyse gewichtet die Datensätze bei der Ermittlung der Beziehungsstärken entsprechend.

Festlegen von Erstellungsoptionen für die Gruppenanalyse

Auf der Registerkarte “Erstellungsoptionen” des Fensters für den Gruppenanalyse-Quellknoten können Sie Optionen für die Ermittlung von Gruppen im Netzwerk definieren.

Einstellungen der Gruppenanalyse

Die Einstellungen der Gruppenanalyse wirken sich auf die Größe und die relative Stärke der gebildeten Gruppen aus.

Schwellenwert Abdeckung. Dient zur Festlegung des Anteils der stärksten Netzwerkbeziehungen, der in der Analyse verwendet werden soll. Ein Abdeckungsschwellenwert von 0,2 beispielsweise gibt an, dass nur die Beziehungen, deren Gewichte zu den stärksten 20 % aller Gewichte gehören, verwendet werden. Durch diesen Parameter, der von 0 bis 1 reicht, kann die Analyse auf die stärksten Beziehungen konzentriert werden, die innerhalb von Gruppen vorkommen sollten. Höhere Schwellenwerte ergeben Gruppen mit schwächeren Beziehungen zwischen ihren Mitgliedern. Für weitere Informationen siehe Thema [Partitionierung in Gruppen](#) auf S. 12.

Mindestgröße der Gruppe. Dient zur Angabe einer Untergrenze für die Gruppengröße. Es werden keine Gruppen ausgegeben, die kleiner sind als dieser Wert.

Maximale Größe der Gruppe. Dient zur Angabe einer Obergrenze für die Gruppengröße. Gruppen, die größer sind als dieser Wert, werden in kleinere Gruppen unterteilt.

Übersichtsstatistiken berechnen und anzeigen. Wenn diese Option ausgewählt ist, leitet der Knoten nicht nur die Ausgabe der Leistungskennzahlen ab, sondern berechnet auch Übersichtsstatistiken für die Analyse und zeigt sie an. Die Berechnung dieser Statistiken kann bei großen Netzwerken die Leistungsfähigkeit des Knotens beeinträchtigen.

Anzeige von Statistiken zur Gruppenanalyse

Auf der Registerkarte “Analyse” des Fensters für den Gruppenanalyse-Quellknoten wird eine überblicksartige Zusammenfassung der im Netzwerk ermittelten Gruppen angezeigt. In der Tabelle “Gruppenanalyse – Übersichtsstatistiken” werden die verfügbaren Übersichtsstatistiken für die Gruppen aufgeführt.

Tabelle 2-2
Gruppenanalyse – Übersichtsstatistiken

Statistik	Beschreibung
Gesamtzahl der Knoten in Gruppen	Anzahl der in den ermittelten Gruppen enthaltenen Knoten.
Gesamtzahl der Verknüpfungen in Gruppen	Anzahl der in den ermittelten Gruppen enthaltenen Verknüpfungen.
Gesamtzahl der Gruppen	Anzahl der im Netzwerk ermittelten Gruppen.
Mittlere Gruppengröße	Durchschnittliche Anzahl an Knoten in einer Gruppe.
Mittlere Gruppendichte	Durchschnittlicher Anteil direkter Beziehungen zwischen Knoten in einer Gruppe. Für weitere Informationen siehe Thema Netzwerkdichte in Kapitel 1 auf S. 4.
Mittlerer Anteil an Kernmitgliedern	Durchschnittlicher Anteil an Knoten in einer Gruppe, bei denen es sich um Kernknoten für die Gruppe handelt. Für weitere Informationen siehe Thema Partitionierung in Gruppen auf S. 12.
Mittlere Dichte der Kerngruppe	Durchschnittlicher Anteil direkter Beziehungen zwischen Kernknoten in einer Gruppe.
Mittlerer In-Degree	Durchschnittliche Anzahl eingehender Verknüpfungen. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.
Mittlerer Out-Degree	Durchschnittliche Anzahl ausgehender Verknüpfungen. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.

Wählen Sie eine bestimmte Statistik in der Tabelle “Übersichtsstatistiken” aus, um die Verteilung der Werte, die Standardabweichung und die Schiefe der Statistik anzuzeigen.

Die Überprüfung dieser Statistiken erleichtert die Festlegung der Einstellungen für die Gruppenanalyse. Idealerweise sollten die Ergebnisse relativ kleine Gruppen mit hohen Dichtewerten zeigen. Wenn beispielsweise einige sehr große Gruppen vorhanden sind, sollten Sie vielleicht die maximale Gruppengröße verringern, um sie in kleinere Gruppen aufzuteilen. Umgekehrt gilt: Bei kleinen Werten für die Gruppendichte sollte eventuell der Schwellenwert für die Abdeckung verringert werden, um die Untersuchung auf stärkere Beziehungen einzugrenzen.

Ausgabe für die Gruppenanalyse

Der Gruppenanalyseknoten generiert verschiedene Felder, die die Gruppen und die einzelnen Personen in den Gruppen beschreiben. Sie können bestehende Modelle und Daten mit diesen Leistungskennzahlen erweitern, um die von Ihren Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Gruppeneinflüsse ergänzen.

In der Tabelle “Leistungskennzahlen für Gruppen” werden die Felder für die Leistungskennzahlen für die in der Analyse ermittelten Gruppen aufgeführt.

Tabelle 2-3
Leistungskennzahlen für Gruppen

Feld	Beschreibung
GAG_GroupNumber	Eindeutige Kennung für eine Gruppe.
GAG_Size	Anzahl der Einzelpersonen in einer Gruppe.
GAG_Density	Anteil direkter Beziehungen zwischen Einzelpersonen in einer Gruppe. Für weitere Informationen siehe Thema Netzwerkdichte in Kapitel 1 auf S. 4.
GAG_KernelDensity	Anteil direkter Beziehungen zwischen zum Kern gehörenden Einzelpersonen in einer Gruppe.
GAG_CoreNodesFraction	Anteil an Einzelpersonen in einer Gruppe, die zum Kern der Gruppe gehören. Für weitere Informationen siehe Thema Partitionierung in Gruppen auf S. 12.
GAG_MaxRankType1	Maximaler Autoritäts-Score für ein beliebiges Gruppenmitglied. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAG_MinRankType1	Minimaler Autoritäts-Score für ein beliebiges Gruppenmitglied.
GAG_MaxMinRankRatioType1	Verhältnis zwischen größtem und kleinstem Autoritäts-Score. Dieser Wert gibt die Autoritätsstärke des Gruppenführers an.
GAG_MaxRankType2	Maximaler Disseminations-Score für ein beliebiges Gruppenmitglied. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAG_MinRankType2	Minimaler Disseminations-Score für ein beliebiges Gruppenmitglied.
GAG_MaxMinRankRatioType2	Verhältnis zwischen größtem und kleinstem Disseminations-Score. Dieser Wert gibt die Disseminationsstärke des Gruppenführers an.

In der Tabelle “Leistungskennzahlen für Einzelpersonen” werden die Felder für die Leistungskennzahlen für die einzelnen Personen im Netzwerk aufgeführt.

Tabelle 2-4
Leistungskennzahlen für Einzelpersonen

Feld	Beschreibung
GAI_NodeNumber	Eindeutige Kennung für eine Einzelperson.
GAI_CoreNode	Kennung, die angibt, ob die Einzelperson zum Kern einer Gruppe gehört oder nicht. Für weitere Informationen siehe Thema Partitionierung in Gruppen auf S. 12.
GAI_RankType1	Autoritäts-Score für die Einzelperson. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAI_RankOrderType1	Rangordnung in der Gruppe auf der Grundlage der Autoritäts-Scores.
GAI_RankType2	Disseminations-Score für die Einzelperson. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAI_RankOrderType2	Rangordnung in der Gruppe auf der Grundlage der Disseminations-Scores.
GAI_InDegree	Anzahl an Beziehungen, bei denen die Einzelperson das Ziel der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.

Feld	Beschreibung
GAI_OutDegree	Anzahl an Beziehungen, bei denen die Einzelperson die Quelle der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.
GAI_GroupLeaderType1	Gibt an, ob der Knoten ein Autoritätsführer ist, dessen Führungs-Score aus eingehenden Links abgeleitet ist. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAI_GroupLeaderConfidenceType1	Die Konfidenz, dass es sich bei dem Knoten um einen Autoritätsführer handelt.
GAI_GroupLeaderType2	Gibt an, ob der Knoten ein Disseminationsführer ist, dessen Führungs-Score aus ausgehenden Links abgeleitet ist. Für weitere Informationen siehe Thema Beschreiben von Gruppen und Gruppenmitgliedern auf S. 13.
GAI_GroupLeaderConfidenceType2	Die Konfidenz, dass es sich bei dem Knoten um einen Disseminationsführer handelt.

Diffusionsanalyse

Diffusionsanalyse – Übersicht

Durch die Diffusionsanalyse werden Einzelpersonen ermittelt, die am meisten durch andere Einzelpersonen in einem sozialen Netzwerk beeinflusst werden. Dabei wird der Effekt als **diffundierte Energie** quantifiziert. Bei diesem Prozess wird ein Ansatz für die Aktivierungsausbreitung verfolgt, bei dem sich ein Effekt iterativ von Netzwerkknoten auf ihre unmittelbaren Nachbarn ausbreitet, wobei er bei der Bewegung von Knoten zu Knoten kleiner wird (Dasgupta, Singh, Viswanathan, Chakraborty, Mukherjea, Nanavati, als auch Joshi, 2008). Wenn ein Knoten Energie erhält, wird er aktiviert und überträgt einen Teil dieser Energie an alle Nachbarn, die Ziele direkter Beziehungen mit dem betreffenden Knoten sind.

Der **Ausbreitungsfaktor** definiert den Anteil der von einem aktivierten Knoten übertragenen Energie. Die restliche Energie verbleibt im Knoten. Alle Knoten, die diese Energie erhalten, übertragen wiederum denselben Anteil an ihre Nachbarn, was zu einem Abklingvorgang bei der Energieübertragung führt. Bei großen Ausbreitungsfaktoren wird mehr Energie gesendet, wodurch die Energie Knoten erreichen kann, die weiter von den ursprünglich aktivierten Knoten entfernt liegen, bevor der Prozess vollständig abklingt. Kleine Ausbreitungsfaktoren führen zu rasch abklingenden Diffusionsprozessen, bei denen die übertragene Energie relativ nahe an den ursprünglichen Knoten bleibt.

Die Gesamtmenge der von einem aktivierten Knoten diffundierten Energie wird auf alle Knoten verteilt, die Ziele gerichteter Beziehungen mit dem Knoten sind. Welche Menge, die einzelnen Knoten erhalten, hängt von der Stärke der Beziehung mit dem aktivierten Knoten ab. Der an einen bestimmten Knoten gesendete Bruchteil ist gleich dem Beziehungsgewicht dividiert durch die Gesamtsumme der Gewichte für alle Beziehungen, in denen der aktivierte Knoten die Quelle ist. Folglich erhalten Nachbarn, die Beziehungen mit höheren relativen Gewichten aufweisen, mehr Energie als Nachbarn mit geringeren relativen Gewichten.

Der Diffusionsprozess endet, wenn eine der folgenden Bedingungen gegeben ist:

- Die aktivierten Knoten sind nicht die Quelle einer gerichteten Beziehung.
- Die übertragene Energiemenge liegt unter dem **Schwellenwert Genauigkeit**, einem Grenzwert für die zu übertragende Energie, die erforderlich ist, damit der Prozess fortgesetzt wird.
- Die Anzahl der Iterationen erreicht eine bestimmte Obergrenze.

Nach Abschluss der Diffusion weisen die Knoten mit der größten Menge an diffundierter Energie die höchste Empfindlichkeit für den Effekt auf, der den Prozess einleitete. Wenn der Prozess beispielsweise bei Knoten beginnt, die abwandern, besteht bei den Knoten mit der höchsten Energie das höchste Risiko für eine Abwanderung. Sie können diesen Knoten spezielle Aufmerksamkeit zukommen lassen, um ihre Abwanderung zu vermeiden.

Beispiel für Diffusionsanalyse

Betrachten Sie das Netzwerk in Kapitel 1 auf S. 3. Das Netzwerk besteht aus sieben Knoten, die direkte Beziehungen unterschiedlicher Stärke zueinander aufweisen.

Die Tabelle "Beispiel für Diffusionsprozess" zeigt einen Diffusionsprozess im Netzwerk mit einem Ausbreitungsfaktor von 0,80. Zunächst enthält Knoten A die gesamte Energie, der willkürlich der Wert 1,00 zugewiesen wurde. In Schritt 1 wird dieser Knoten aktiviert, wodurch 80 % seiner Energie auf seine drei Nachbarn verbreitet werden, die Ziele von Beziehungen sind, während 20 % beim ursprünglichen Knoten verbleiben. Die Beziehungen haben ein Gesamtgewicht von 100. Die Beziehung mit Knoten D macht die Hälfte dieses Gesamtwerts aus, dieser Knoten enthält somit die Hälfte der diffundierten Energie, also 0,40. Die Beziehung mit Knoten B macht 20 % des Gesamtgewichts aus, was den Wert 0,16 für die diffundierte Energie ergibt. Knoten C erhält die restliche Energie, 0,24, das sind 30 % der aus Knoten A stammenden Energiemenge.

Tabelle 3-1
Beispiel für Diffusionsprozess

Schritt	Knoten A	Knoten B	Knoten C	Knoten D	Knoten E	Knoten F	Knoten 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

Nachdem die Knoten B, C und D in Schritt 1 Energie erhalten haben, werden sie in Schritt 2 aktiviert und diffundieren Energie an ihre Zielnachbarn (sofern vorhanden). Knoten B verfügt über keine Zielknoten und diffundiert daher keine Energie. Knoten C dagegen verfügt über ein Ziel und diffundiert 80 % seiner Energie an den betreffenden Knoten, Knoten G. Knoten D schließlich verbreitet 80 % seiner Energie an die Knoten E und F. Dabei enthält F das Dreifache von E, da er eine stärkere Beziehung mit D aufweist.

In Schritt 3 werden die Knoten E, F und G aktiviert. Knoten E und G haben keine Zielknoten und diffundieren daher keine Energie. Knoten F dagegen diffundiert 80 % seiner Energie an seinen einzigen Zielknoten, G. Dieser Knoten weist nun einen Gesamtenergiewert von 0,38 auf, der sich aus der von den Knoten C und F diffundierten Energiemengen zusammensetzt.

Nun wird Knoten G aktiviert. Er verfügt jedoch nicht über Zielknoten für die Diffusion, weshalb der Prozess endet. Die von Knoten A herrührende Energie hat sich über das Netzwerk ausgebreitet, wobei Knoten G die größte Energiemenge erhielt. Wenn die in das Netzwerk eingeführte Energie für Abwanderung steht, wäre Knoten G am meisten von einer Abwanderung von Knoten A betroffen.

Diffusionsanalyseknoten

Der Diffusionsanalyseknoten, der über die Palette "Datenquellen" verfügbar ist, verbreitet einen Effekt aus einer angegebenen Menge von Einzelpersonen in einem sozialen Netzwerk. Dabei werden die Netzwerkbeziehungen verwendet, um die Personen zu ermitteln, die am stärksten von dem Effekt betroffen sind. Wenn es sich bei dem Effekt beispielsweise um Abwanderung handelt,

gibt der Knoten diejenigen Personen an, bei denen die Wahrscheinlichkeit am höchsten ist, dass sie abwandern, wenn bestimmte andere Personen im Netzwerk abwandern. Sie können bestehende Modelle und Daten mit der Knotenausgabe erweitern, um die von diesen Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Diffusionseinflüsse ergänzen.

Anforderungen. Für den Knoten sind zwei Textdateien mit fester Breite erforderlich, die die zu analysierenden Daten enthalten. Die erste Datei definiert das soziale Netzwerk mittels dreier Felder. Ein Feld gibt die Quelle für jede gerichtete Beziehung an, ein Feld gibt das Ziel für jede gerichtete Beziehung an und das dritte Feld gibt eine optionale Stärke für jede Beziehung an. Alle Beziehungen im Netzwerk müssen gerichtet sein. Die zweite Datei enthält eine Liste mit Kennzeichnern, die angeben, von wo der Effekt beginnen soll.

Angabe von Daten für die Diffusionsanalyse

Auf der Registerkarte "Daten" des Fensters für den Diffusionsanalyse-Quellknoten können Sie die Eingabedatei angeben, die die Beziehungen zwischen den Knoten im Netzwerk enthält.

Datei. Geben Sie den Namen der Datei oder des Ordners mit den Kommunikationsdatensätzen ein. Zur Auswahl eines Namens aus dem Dateisystem können Sie einen Dateinamen eingeben oder auf die Schaltfläche mit den Auslassungspunkten (...) klicken. Der Pfad wird angezeigt, sobald Sie einen Namen ausgewählt haben, und der entsprechende Inhalt wird mit Trennzeichen angezeigt. Wenn Sie einen Ordner angeben, werden die Kommunikationsdatensätze aller in dem Ordner befindlichen Dateien für die Analyse aneinandergereiht. Dabei müssen alle Dateien in dem Ordner die gleiche Struktur haben.

Feldnamen aus Datei lesen. Diese standardmäßig ausgewählte Option behandelt die erste Zeile der Datendatei als Namen für die Spalte. Wenn die erste Zeile keine Überschrift ist, deaktivieren Sie diese Option, um jedem Feld automatisch einen generischen Namen zuzuweisen, beispielsweise *Feld1* und *Feld2*.

Einstellungen für Netzwerkdefinition

Die Einstellungen für die Netzwerkdefinition legen die Rollen der Felder fest.

Felder. Verwenden Sie die Pfeilschaltflächen, um Elemente aus dieser Liste manuell den verschiedenen Rollenfeldern zuzuweisen. Die Symbole zeigen das gültige Messniveau für jedes Rollenfeld an. Klicken Sie auf die Schaltfläche *Alle*, damit alle Felder in der Liste ausgewählt werden, oder klicken Sie auf die Schaltfläche für ein individuelles Messniveau, um alle Felder mit diesem Messniveau auszuwählen.

Quelle. Wählen Sie ein Feld als Ursprung der gerichteten Beziehung aus.

Ziel. Wählen Sie ein Feld als Ziel der gerichteten Beziehung aus.

Gewicht. Wählen Sie ein Feld aus, das für das Beziehungsgewicht steht. Bei Kommunikationsdatensätzen könnte das Gewicht einer Anzahl von Anrufmerkmalen entsprechen, die die Beziehungsstärke widerspiegeln, beispielsweise Anrufdauer oder Anruhfrequenz.

Festlegen von Erstellungsoptionen für die Diffusionsanalyse

Auf der Registerkarte “Erstellungsoptionen” des Fensters für den Diffusionsanalyse-Quellknoten können Sie Optionen für den Diffusionsprozess definieren.

Liste der Diffusionsstartwerte

Die Liste der Diffusionsstartwerte gibt die Netzwerkknoten an, von denen aus die Energie diffundieren soll. Bei einer Abwanderungsanalyse beispielsweise gibt diese Liste die Netzwerkknoten an, die bereits abgewandert sind. Alternativ könnte diese Liste Knoten enthalten, bei denen das Abwanderungsrisiko besonders hoch ist, um die Auswirkungen auf das Netzwerk für den Fall zu untersuchen, dass diese Knoten tatsächlich abwandern.

Datei. Geben Sie den Namen der Textdatei an, die die Knotenkennungen enthält. Dabei entspricht jede Kennung einer Zeile in der Datei. Zur Auswahl einer Datei können Sie einen Dateinamen eingeben oder auf die Schaltfläche mit den Auslassungspunkten (...) klicken.

Einstellungen der Diffusionsanalyse

Die Einstellungen für die Diffusionsanalyse legen fest, wie aggressiv die Diffusion ist und wann der Prozess endet.

Ausbreitungsfaktor. Dient zur Angabe des Anteils der Energie, den ein Knoten in jedem Schritt des Diffusionsprozesses überträgt. Mit kleineren Werten wird die Energie auf die Knoten übertragen, die dem Knoten, der die Energie diffundiert, am nächsten liegen. Entsprechend ermöglichen größere Werte, dass sich die Energie auf weiter entfernt liegende Knoten auswirkt.

Maximalzahl der Iterationen. Dient zur Angabe einer Obergrenze für die Anzahl an Diffusionsiterationen. Der Diffusionsprozess endet automatisch, wenn diese Obergrenze erreicht wird.

Schwellenwert Genauigkeit Dient zur Angabe der kleinsten Änderung in der Knotenenergie, die eine Fortsetzung des Diffusionsprozesses rechtfertigt. Wenn die Änderung in der Energie für jeden Knoten kleiner ist als dieser Wert, endet der Diffusionsprozess.

Übersichtsstatistiken berechnen und anzeigen. Wenn diese Option ausgewählt ist, leitet der Knoten nicht nur die Ausgabe der Leistungskennzahlen ab, sondern berechnet auch Übersichtsstatistiken für die Analyse und zeigt sie an. Die Berechnung dieser Statistiken kann bei großen Netzwerken die Leistungsfähigkeit des Knotens beeinträchtigen.

Anzeige von Statistiken zur Diffusionsanalyse

Auf der Registerkarte “Analyse” des Fensters für den Diffusionsanalyse-Quellknoten wird eine überblicksartige Zusammenfassung der Diffusionsergebnisse angezeigt. In der Tabelle “Diffusionsanalyse – Übersichtsstatistiken” werden die verfügbaren Übersichtsstatistiken aufgeführt.

Tabelle 3-2
Diffusionsanalyse – Übersichtsstatistiken

Statistik	Beschreibung
Gesamtzahl der Knoten im Netzwerk	Anzahl der Knoten im Netzwerk.
Gesamtzahl der Zusammenhänge im Netzwerk	Anzahl der Zusammenhänge im Netzwerk.
Gesamtzahl der Diffusionsstartwerte im Netzwerk	Anzahl der als Startwerte für den Diffusionsprozess verwendeten Knoten.
Mittelwert Einfluss	Durchschnittsmenge an diffundierter Energie, die Einzelpersonen zugeordnet ist.
Mittelwert InDegree	Durchschnittliche Anzahl an Beziehungen, bei denen eine Einzelperson das Ziel der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.
Mittelwert OutDegree	Durchschnittliche Anzahl an Beziehungen, bei denen eine Einzelperson die Quelle der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.

Wählen Sie eine bestimmte Statistik in der Tabelle “Übersichtsstatistiken” aus, um die Verteilung der Werte, die Standardabweichung und die Schiefe der Statistik anzuzeigen.

Ausgabe für die Diffusionsanalyse

Der Diffusionsanalyseknoten generiert verschiedene Felder, die die einzelnen Personen im Netzwerk beschreiben. Sie können bestehende Modelle und Daten mit diesen Leistungskennzahlen erweitern, um die von Ihren Modellen generierten Vorhersagen zu verbessern. Sie können beispielsweise die Werte für die Abwanderungsneigung der einzelnen Personen um die Diffusionseinflüsse ergänzen.

In der Tabelle “Diffusionsanalyse – Leistungskennzahlen” werden die Felder für die Leistungskennzahlen für die einzelnen Personen in der Analyse aufgeführt.

Tabelle 3-3
Diffusionsanalyse – Leistungskennzahlen

Feld	Beschreibung
DA_NodeNumber	Eindeutige Kennung für eine Einzelperson.
DA_DiffusedEnergy	Menge an diffundierter Energie, die der Einzelperson zugeordnet ist. Bei der Abwanderungsanalyse weisen höhere Werte auf eine größere Abwanderungsneigung hin an als niedrigere Werte.
DA_InDegree	Anzahl an Beziehungen, bei denen die Einzelperson das Ziel der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.
DA_OutDegree	Anzahl an Beziehungen, bei denen die Einzelperson die Quelle der Beziehung ist. Für weitere Informationen siehe Thema Knotengrad in Kapitel 1 auf S. 5.

Hinweise

Diese Informationen wurden für weltweit angebotene Produkte und Dienstleistungen erarbeitet.

IBM bietet die in diesem Dokument behandelten Produkte, Dienstleistungen oder Merkmale möglicherweise nicht in anderen Ländern an. Informationen zu den derzeit in Ihrem Land erhältlichen Produkten und Dienstleistungen erhalten Sie bei Ihrem zuständigen IBM-Mitarbeiter vor Ort. Mit etwaigen Verweisen auf Produkte, Programme oder Dienste von IBM soll nicht behauptet oder impliziert werden, dass nur das betreffende Produkt oder Programm bzw. der betreffende Dienst von IBM verwendet werden kann. Stattdessen können alle funktional gleichwertigen Produkte, Programme oder Dienste verwendet werden, die keine geistigen Eigentumsrechte von IBM verletzen. Es obliegt jedoch der Verantwortung des Benutzers, die Funktionsweise von Produkten, Programmen oder Diensten von Drittanbietern zu bewerten und zu überprüfen.

IBM verfügt möglicherweise über Patente oder hat Patentanträge gestellt, die sich auf in diesem Dokument beschriebene Inhalte beziehen. Durch die Bereitstellung dieses Dokuments werden Ihnen keinerlei Lizenzen an diesen Patenten gewährt. Lizenzanfragen können schriftlich an folgende Adresse gesendet werden:

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

Der folgende Abschnitt findet in Großbritannien und anderen Ländern keine Anwendung, in denen solche Bestimmungen nicht mit der örtlichen Gesetzgebung vereinbar sind: INTERNATIONAL BUSINESS MACHINES STELLT DIESE VERÖFFENTLICHUNG IN DER VERFÜGBAREN FORM OHNE GARANTIE BEREIT, SEIEN ES AUSDRÜCKLICHE ODER STILLSCHWEIGENDE, EINSCHLIESSLICH JEDOCH NICHT NUR DER GARANTIE BEZÜGLICH DER NICHT-RECHTSVERLETZUNG, DER GÜTE UND DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK. Manche Rechtsprechungen lassen den Ausschluss ausdrücklicher oder implizierter Garantien bei bestimmten Transaktionen nicht zu, sodass die oben genannte Ausschlussklausel möglicherweise nicht für Sie relevant ist.

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler aufweisen. An den hierin enthaltenen Informationen werden regelmäßig Änderungen vorgenommen. Diese Änderungen werden in neuen Ausgaben der Veröffentlichung aufgenommen. IBM kann jederzeit und ohne vorherige Ankündigung Optimierungen und/oder Änderungen an den Produkten und/oder Programmen vornehmen, die in dieser Veröffentlichung beschrieben werden.

Jegliche Verweise auf Drittanbieter-Websites in dieser Information werden nur der Vollständigkeit halber bereitgestellt und dienen nicht als Befürwortung dieser. Das Material auf diesen Websites ist kein Bestandteil des Materials zu diesem IBM-Produkt und die Verwendung erfolgt auf eigene Gefahr.

IBM kann die von Ihnen angegebenen Informationen verwenden oder weitergeben, wie dies angemessen erscheint, ohne Ihnen gegenüber eine Verpflichtung einzugehen.

Lizenznehmer dieses Programms, die Informationen dazu benötigen, wie (i) der Austausch von Informationen zwischen unabhängig erstellten Programmen und anderen Programmen und (ii) die gegenseitige Verwendung dieser ausgetauschten Informationen ermöglicht wird, wenden sich an:

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

Derartige Informationen stehen ggf. in Abhängigkeit von den jeweiligen Geschäftsbedingungen sowie in einigen Fällen der Zahlung einer Gebühr zur Verfügung.

Das in diesem Dokument beschriebene lizenzierte Programm und sämtliche dafür verfügbaren lizenzierten Materialien werden von IBM gemäß dem IBM-Kundenvertrag, den Internationalen Nutzungsbedingungen für Programmpakete der IBM oder einer anderen zwischen uns getroffenen Vereinbarung bereitgestellt.

Diese Informationen enthalten Beispiele zu Daten und Berichten, die im täglichen Geschäftsbetrieb Verwendung finden. Um diese so vollständig wie möglich zu illustrieren, umfassen die Beispiele Namen von Personen, Unternehmen, Marken und Produkten. Alle diese Namen sind fiktiv und jegliche Ähnlichkeit mit Namen und Adressen realer Unternehmen ist rein zufällig.

Marken

IBM, das IBM-Logo, ibm.com und SPSS sind Marken der IBM Corporation und in vielen Ländern weltweit registriert. Eine aktuelle Liste der IBM-Marken finden Sie im Internet unter <http://www.ibm.com/legal/copytrade.shtml>.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA, anderen Ländern oder beidem.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA, anderen Ländern oder beidem.

UNIX ist eine eingetragene Marke der The Open Group in den USA und anderen Ländern.

Andere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein.

Glossar

Ausbreitungsfaktor. Der Anteil an Energie, den ein Knoten in einem Schritt des Diffusionsprozesses an seine Nachbarn überträgt.

Autoritäts-Score. Maßzahl, die angibt, in welchem Umfang die anderen Mitglieder einer Gruppe eine Verbindung zu einer Einzelperson herstellen. Autoritäts-Scores verknüpfen die Bedeutung einer Einzelperson mit der Anzahl der Beziehungen, die bei dieser Person enden.

Autoritätsführer. Die Einzelperson in einer Gruppe, die den höchsten Autoritäts-Score aufweist.

Bewertete Beziehung. Eine Verbindung zwischen zwei Knoten, denen ein Gewicht zugeordnet ist, das die Stärke der Beziehung angibt.

Dichotome Beziehung. Eine Verbindung zwischen zwei Knoten, die nur einen von zwei Werten annehmen kann, die für das Vorhandensein bzw. Fehlen der Verbindung stehen.

Dichte. Eine Maßzahl der Netzwerkkohäsion, die als Anzahl der beobachteten Beziehungen dividiert durch die Anzahl der möglichen Beziehungen definiert ist.

Diffundierte Energie. Effekt, der von bestimmten Knoten in einem Netzwerk ausgeht und sich über das gesamte Netzwerk ausbreitet, wobei seine Größe mit zunehmender Entfernung von den Ursprungsknoten abnimmt.

Disseminationsführer. Die Einzelperson in einer Gruppe, die den höchsten Disseminations-Score aufweist.

Disseminations-Score. Maßzahl, die angibt, in welchem Umfang eine Einzelperson eine Verbindung zu anderen Mitglieder einer Gruppe herstellt. Disseminations-Scores verknüpfen die Bedeutung einer Einzelperson mit der Anzahl der Beziehungen, die von dieser Person ausgehen.

Gerichtete Beziehung. Eine Verbindung zwischen zwei Knoten, die von einem Knoten (der Quelle) ausgeht und beim anderen Knoten (dem Ziel) endet.

Grad. Maßzahl der Knotenaktivität in einem Netzwerk, definiert als Gesamtzahl der Beziehungen, an denen der Knoten beteiligt ist.

Indegree. Maßzahl des Prestiges für einen Knoten in einem Netzwerk, das aus gerichteten Beziehungen besteht, definiert als Gesamtzahl der Beziehungen, in denen der Knoten das Ziel ist.

Kerngruppe. Eine Gruppe von Knoten, die verbleiben, nachdem schwächere Beziehungen aus einem Netzwerk weggelassen und Größenbeschränkungen auf die Gruppe angewendet wurden.

Outdegree. Maßzahl der Zentralität für einen Knoten in einem Netzwerk, das aus gerichteten Beziehungen besteht, definiert als Gesamtzahl der Beziehungen, in denen der Knoten die Quelle ist.

Schwellenwert Abdeckung. Der Anteil der stärksten Beziehungen, der für die Gruppenidentifizierung verwendet wird.

Schwellenwert Genauigkeit. Ein Stoppkriterium für einen Diffusionsprozess, das der kleinsten Änderung in der Knotenenergie entspricht, die die Fortsetzung der Diffusion rechtfertigt.

Soziogramm. Eine visuelle Darstellung eines sozialen Netzwerks, in der mit Linien verbundene Punkte für Personen stehen, die eine Beziehung zueinander haben.

Ungerichtete Beziehung. Eine Verknüpfung zwischen Knoten.

Bibliografie

Cover, T. M., als auch 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, als auch 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, als auch 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., als auch K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Index

- Analyse (Registerkarte), 7
 - Diffusionsanalyse, 22
 - Gruppenanalyse, 16
- Analysieren von Daten, 8
- Anmerkungen (Registerkarte), 7
- Anzeigen einer Ausgabevorschau, 7
- Ausbreitungsfaktor, 19, 22
- Ausführen von Streams
 - Verwendung von IBM SPSS Modeler Server Social Network Analysis, 7
- Autoritäts-Scores, 13, 17
- Autoritätsführer, 13, 18

- Bewertete Beziehungen, 2
- Beziehung
 - Bewertet, 2
 - Richtung, 2

- Daten (Registerkarte), 7
 - Diffusionsanalyse, 21
 - Gruppenanalyse, 14
- Dichotome Beziehungen, 2
- Dichte, 4, 16–17
- Diffundierte Energie, 23
- Diffusionsanalyse, 6, 20
 - Anforderungen, 21
 - Ausbreitungsfaktor, 19, 22
 - Optionen, 21
 - Schwellenwert Genauigkeit, 19, 22
 - Startwerte, 22
- Disseminations-Scores, 13, 17
- Disseminationsführer, 13, 18

- Eigenschaften
 - Skripts, 9
- Erstellungsoptionen (Registerkarte), 7
 - Diffusionsanalyse, 22
 - Gruppenanalyse, 15

- Filter (Registerkarte), 7
- Filtern der Ausgaben, 7

- GAG_GroupNumber, 17
- Gerichtete Beziehungen, 2
- Gewichtungsfelder, 9, 15, 21
- Grad, 5
- Gruppenanalyse, 6
 - Anforderungen, 14
 - Gruppengrößen, 13, 15
 - Schwellenwert Abdeckung, 12, 15
- Gruppengrößen, 16–17
 - Grenzwerte, 13, 15

- In-Degree, 16
- Indegree, 5, 17, 23

- Kerndichte, 16–17
- Kerngruppen, 13
- Kernknoten, 16–17
- Kommunikationsdatensätze, 9

- Leistungskennzahlen, 16, 23
- Liste der Diffusionsstartwerte, 22
- Löschen von Ergebnissen, 8

- Marken, 25
- Messniveaus, 7
- Mittelwert Einfluss, 23
- Mittelwert Indegree, 23
- Mittelwert Outdegree, 23
- Mittlere Gruppendichte, 16
- Mittlere Gruppengröße, 16
- Mittlere Kerngruppendichte, 16
- Mittlerer Anteil an Kernmitgliedern, 16
- Mittlerer In-Degree, 16
- Mittlerer Out-Degree, 16

- Out-Degree, 16
- Outdegree, 5, 18, 23

- Quellenfelder, 9, 15, 21

- Rechtliche Hinweise, 24

- Schwellenwert Abdeckung, 12, 15
- Schwellenwert Genauigkeit, 19, 22
- Skripts
 - Eigenschaften, 9
 - Startwerte
 - für Diffusionsanalyse, 22

- Typen (Registerkarte), 7

- Ungerichtete Beziehungen, 2

- Zielfelder, 9, 15, 21