

*IBM SPSS Modeler Social Network
Analysis 18.0 — podręcznik
użytkownika*

IBM

Uwaga

Przed skorzystaniem z niniejszych informacji oraz produktu, którego one dotyczą, należy zapoznać się z informacjami zamieszczonymi w sekcji "Uwagi" na stronie 19.

Informacje o produkcie

Niniejsze wydanie publikacji dotyczy wersji 18, wydania 0, modyfikacji 0 produktu IBM SPSS Modeler oraz wszystkich następnych wersji i modyfikacji do czasu, aż w kolejnym wydaniu publikacji zostanie zawarta informacja o stosownej zmianie.

Spis treści

Przedmowa v

Informacje o programie IBM Business Analytics v

Wsparcie techniczne v

Rozdział 1. IBM SPSS Modeler Social Network Analysis. 1

O programie IBM SPSS Modeler Social Network Analysis 1

Informacje o analizie sieci społecznościowych. 1

 Pokazywanie sieci 2

 Opisywanie sieci 3

Węzły IBM SPSS Modeler Social Network Analysis 5

 Zakładki węzłów 5

 Przeglądanie wyników 5

 Analizowanie danych 6

Zastosowania 6

Struktura danych 7

Właściwości skryptów 7

Rozdział 2. Analiza grup 9

Przegląd analizy grup 9

 Określenie podobieństw społecznych 9

 Dzielenie na grupy 10

 Opisywanie grup i członków grup 10

Węzeł Analizy grup 11

 Określenie danych do analizy grup 11

 Ustawianie opcji kompilacji dla analizy grup 12

 Przeglądanie statystyki analizy grup 12

 Dane wyjściowe do analizy grup 13

Rozdział 3. Analiza przenikania 15

Przegląd analizy przenikania. 15

 Przykład analizy przenikania 15

Węzeł Analizy przenikania 16

 Określenie danych do analizy przenikania. 16

 Ustawianie opcji kompilacji dla analizy przenikania. 17

 Przeglądanie statystyki analizy przenikania 17

 Dane wyjściowe dla analizy przenikania 18

Uwagi 19

Znaki towarowe 20

Warunki dotyczące dokumentacji produktu 21

Indeks 25

Przedmowa

IBM® SPSS Modeler Social Network Analysis przekształca informacje dotyczące relacji między ludźmi w pola opisujące rolę danej osoby w sieci społecznościowej, pozwalając na uwzględnienie informacji społecznych w modelach predykcyjnych. Niniejsza instrukcja opisuje sposób wykorzystania węzłów IBM SPSS Modeler Social Network Analysis w środowisku IBM SPSS Modeler, pozwalając użytkownikowi na ujęcie węzłów w jego strumieniach. Po połączeniu danych wyjściowych węzła z polami reprezentującymi mierniki osób, powstaje bardziej kompletny profil tych osób.

Informacje o programie IBM Business Analytics

Oprogramowanie IBM Business Analytics dostarcza kompletne, spójne i dokładne informacje, na których mogą polegać osoby decyzyjne chcąc polepszyć wyniki biznesowe. Wszechstronne portfolio obejmujące moduły analiza biznesowa, analiza prognostyczna, zarządzanie wynikami i strategiami finansowymi oraz aplikacje analityczne zapewnia jasny, natychmiastowy i pozwalający na podjęcie działań wgląd w bieżące wyniki oraz daje możliwość przewidywania przyszłych wyników. W połączeniu z licznymi rozwiązaniami branżowymi, sprawdzonymi praktykami i profesjonalnymi usługami, organizacje o różnych rozmiarach mogą wspomagać najwyższą produktywność, w sposób pewny zautomatyzować decyzje i uzyskać lepsze wyniki.

Oprogramowanie IBM SPSS Predictive Analytics, będąc częścią tego portfolio, wspomaga organizacje w zakresie przewidywania przyszłych zdarzeń oraz proaktywnie wpływać na ten wgląd w celu wspomaganie lepszych wyników finansowych. Klienci komercyjni, rządowi i uczelnie na całym świecie polegają na technologii IBM SPSS, zapewniającej przewagę konkurencyjną przyciągającą, zatrzymującą i rozwijającą klientów, zmniejszając nieuczciwość i zmniejszając ryzyko. Wdrażając oprogramowanie IBM SPSS do swojej codziennej działalności, organizacje stają się przewidującymi przedsiębiorstwami, zdolnymi do zarządzania i automatyzacji decyzji w celu realizacji celów biznesowych i osiągnięcia mierzalnej przewagi konkurencyjnej. W celu uzyskania dalszych informacji lub skontaktowania się z przedstawicielem, proszę wejść na stronę <http://www.ibm.com/spss>.

Wsparcie techniczne

Wsparcie techniczne jest dostępne w celu zapewnienia klientom obsługi technicznej. Klienci mogą się kontaktować z działem Wsparcia technicznego w celu uzyskania pomocy dotyczącej korzystania z IBM Corp. produktów lub pomocy w instalacji dla jednego z obsługiwanych środowisk sprzętowych. Aby skontaktować się z działem Wsparcia technicznego, wejdź na stronę internetową IBM Corp. pod adresem <http://www.ibm.com/support>. W przypadku prośby o pomoc, należy przygotować swoje dane identyfikacyjne, dane swojej organizacji, a także dane dotyczące usług wsparcia.

Rozdział 1. IBM SPSS Modeler Social Network Analysis

O programie IBM SPSS Modeler Social Network Analysis

Wiele podejść do modelowania zachowania koncentruje się na pojedynczych osobach. Wykorzystują one różne dane na temat pojedynczych osób w celu generowania modelu, który do przewidywania zachowania wykorzystuje kluczowe wskaźniki zachowania. Jeśli dla jakiejś osoby dostępne są wartości kluczowych wskaźników, które są powiązane z wystąpieniem zachowania, osoba ta może stać się celem szczególnej uwagi mającej na celu zapobieganie temu zachowaniu.

Rozważ podejścia do modelowania odchodzenia klientów, w których klient kończy relację z firmą. Koszt zatrzymania klientów jest znacznie niższy od kosztu zastąpienia ich nowymi, co powoduje, że zdolność identyfikacji klientów, dla których istnieje ryzyko odejścia, jest podstawową kwestią. Analityk często wykorzystuje wiele Kluczowych wskaźników wydajności do opisywania klientów, uwzględniając informacje demograficzne, oraz wzorce ostatnich rozmów z klientami dla każdego klienta z osobna. Modele predykcyjne na podstawie tych pól wykorzystują zmiany wzorców rozmów telefonicznych z klientami, które są zgodne z wzorcami rozmów telefonicznych z klientami, którzy odeszli w przeszłości, aby w przyszłości zidentyfikować ludzi wykazujących zwiększone ryzyko odejścia. Klienci zidentyfikowani jako należący do grupy ryzyka otrzymują dodatkowe wsparcie lub oferty wsparcia, co ma prowadzić do ich zatrzymania.

Metody te nie uwzględniają informacji społecznych, które w znacznym stopniu mogą wpływać na zachowanie klienta. Informacje o firmie oraz o tym, co robią inni ludzie, przepływają dzięki relacjom, aby ostatecznie wpływać na ludzi. W wyniku tego relacje z innymi ludźmi pozwalają tym ludziom wpływać na decyzje i działania danej osoby. Analizy zawierające tylko pojedyncze mierniki pomijają istotne czynniki, które umożliwiają przewidywanie.

IBM SPSS Modeler Social Network Analysis zajmuje się tym problemem przetwarzając informacje o relacji w dodatkowe pola, które można uwzględnić w modelach. Te podzielone kluczowe wskaźniki wydajności mierzą właściwości społeczne dla pojedynczych osób. Połączenie tych cech społecznych z miernikami bazującymi na poszczególnych osobach zapewniają lepszy przegląd poszczególnych osób i w efekcie mogą ulepszyć dokładność przewidywania Twoich modeli.

IBM SPSS Modeler Social Network Analysis składa się z dwóch podstawowych składników:

- Węzły IBM SPSS Modeler Social Network Analysis dodane do środowiska IBM SPSS Modeler, które umożliwiają dołączenie technik analizy społecznej do strumieni.
- IBM SPSS Modeler Server Social Network Analysis, co powoduje dodanie przetwarzania specyfikacji węzła do IBM SPSS Modeler Server. IBM SPSS Modeler Server Social Network Analysis efektywnie przetwarza znaczne ilości danych sieciowych, które mogą zawierać miliony osób i relacji, we względnie małą ilość pól do dalszej analizy.

Na przykład, IBM SPSS Modeler Social Network Analysis identyfikuje te osoby w sieci, których w największym stopniu dotyczy określony sposób odchodzenia. Ponadto możliwe jest odkrycie grup ludzi, których dotyczy zwiększone ryzyko odejścia. Poprzez wdrożenie Kluczowych wskaźników wydajności dla tych efektów w Twoich modelach możesz polepszyć ich ogólną wydajność.

Informacje o analizie sieci społecznościowych

Sieć społecznościowa składa się z zestawu pojedynczych osób i relacji między nimi. Analiza sieci społecznościowych bada te relacje w celu opisanego pojedynczych osób i grup jako części struktury społecznej. Pojedyncze osoby współdziałają ze sobą, a wzorce interakcji dają wgląd w to, jakie to są osoby. Relacje umożliwiają przepływ informacji w sieci, pozwalając wpływać pewnym osobom na inne. Znaczenie informacji o relacjach oddziela analizę sieci społecznościowych od innych podejść. Zamiast koncentrować się na każdej osobie z osobna, przedmiotem badań jest para składająca się z dwóch osób i ich relacji.

Relacje w sieci można sklasyfikować jako kierunkowe lub niekierunkowe. W *relacji kierunkowej* jedną osobę identyfikuje się jako inicjatora, lub źródło, relacji a drugą identyfikuje się jako odbiorcę, lub cel. Na przykład rozmowa telefoniczna jest relacją kierunkową, w której jedna osoba dzwoni do drugiej. Z kolei zdefiniowanie ról źródła i celu dla *relacji niekierunkowych* jest niemożliwe. W takim przypadku obie strony w równym stopniu biorą udział w relacji. Rozmowa dwóch osób jest przykładem relacji niekierunkowej

Inną właściwością, która rozróżnia relacje jest to, czy relacja jest dychotomiczna, czy jest to relacja z wartościowaniem. Jedyną informacją, która jest dostępna w *relacji dychotomicznej* jest to, czy między dwoma osobami występuje relacja. Dla każdej pary w sieci relacja albo występuje, albo nie. Z drugiej strony *relacja z wartościowaniem* zawiera wagę wskazującą siłę relacji. Wagi pozwalają na porównanie ze sobą poszczególnych relacji.

Tabela „Typy relacji” zawiera przykład klasyfikacji krzyżowej relacji według kierunku i skali. W relacjach kierunkowych, *Jan* jest źródłem relacji, a *Maria* jest celem. W relacjach niekierunkowych brak jest wskazania, kto rozpoczął relację. Relacje z wartościowaniem wykorzystują długość konwersacji jako wagę relacji, podczas gdy relacje dychotomiczne albo wystąpiły, albo nie.

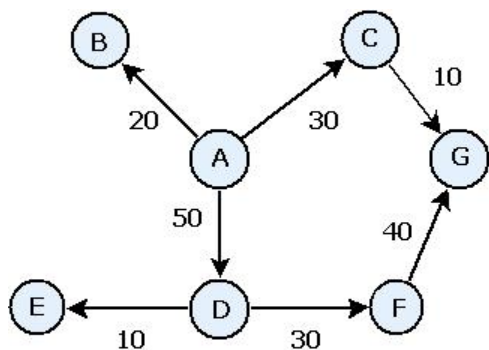
Tabela 1. Typy relacji.

Kierunek	Skala	Przykład
Niekierunkowa	Dychotomiczna	Jan i Maria rozmawiali ze sobą
Niekierunkowa	Z wartościowaniem	Jan i Maria rozmawiali ze sobą przez 20 minut
Kierunkowa	Dychotomiczna	Jan zadzwonił do Marii
Kierunkowa	Z wartościowaniem	Jan przeprowadził z Marią 20-minutową rozmowę telefoniczną

W celu uzyskania dalszych informacji w zakresie analizy sieci społecznościowych należy sięgnąć po jedną z wyczerpujących książek na ten temat ¹.

Pokazywanie sieci

Sieć społecznościową zazwyczaj ilustruje się za pomocą *socjogramu* ². W wizualizacji tego typu poszczególne osoby są przedstawiane za pomocą punktów lub węzłów w przestrzeni. Linie lub krawędzie łączące punkty przedstawiają relacje między poszczególnymi osobami. Jeśli relacje są kierunkowe, krawędzie zawierają strzałkę wskazującą kierunek. Jeśli relacje zawierają wagi, etykiety krawędzi oznaczają wartości. Poniższy wykres przedstawia sieć składającą się z siedmiu osób.



Rysunek 1. Przykład sieci społecznościowej

Załóżmy, że sieć przedstawia rozmowy telefoniczne przeprowadzone przez osoby, gdzie wagi wskazują długość rozmów. W takim przypadku Osoba A zadzwoniła do trzech osób, rozmawiając przez większość czasu z osobą D.

1. Wasserman, S. i K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

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

Ta sieć jest znacznie mniejsza od tych, które spotyka się w praktyce. Jednak pojęcia ilustrowane za pomocą zwykłych socjogramów odnoszą się do sieci o dowolnych rozmiarach i stopniu skomplikowania.

Opisywanie sieci

Informacje o sieciach, grupach i poszczególnych osobach trzeba wydobyć w taki sposób, aby przybrały postać cech opisowych, pozwalających na porównania krzyżowe i ujęcie ich w modelach predykcyjnych. Z sieci należy pozyskać skończony zestaw kluczowych wskaźników wydajności, które nadają się do analizy. Na przykład użytkownik może chcieć porównać sieci lub grupy węzłów w ramach sieci lub z osobna. Alternatywnie użytkownik może chcieć porównać poszczególne osoby w sieci lub ustalić osoby o największym znaczeniu.

Dwoma miernikami, których często się używa do opisywania sieci społecznościowych, są **gęstość** i **stopień**. Oba parametry odzwierciedlają spójność, ale pierwszy z nich koncentruje się na całej sieci lub podgrupach sieci, podczas gdy drugi charakteryzuje poszczególne osoby w sieci.

Gęstość sieci

Dla każdego zestawu węzłów w sieci istnieje skończona liczba możliwych relacji. Każdy węzeł może służyć jako źródło lub cel relacji z każdym innym węzłem. Załóżmy, że mamy sieć składającą się z trzech węzłów A, B i C. Poniższa tabela przedstawia wszystkie możliwe ukierunkowane relacje między tymi węzłami.

Tabela 2. Możliwe ukierunkowane relacje dla trzech węzłów.

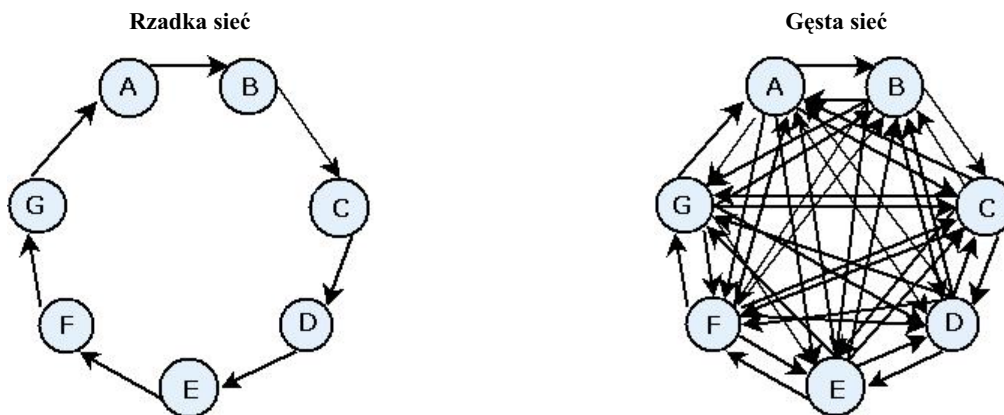
Źródło	Cel
A	B
A	C
B	A
B	C
C	A
C	B

Każde powiązanie jest źródłem relacji z dwoma innymi węzłami. Jednak w praktyce wystąpienie wszystkich możliwych relacji jest niemożliwe. Niektóre węzły mogą nie mieć żadnych bezpośrednich relacji z innymi węzłami. Ponadto niektóre skierowane relacje mogą nie być odwzajemnione.

Parametr *gęstość* przedstawia proporcje możliwych relacji w sieci, które rzeczywiście występują. Wartość ta mieści się w przedziale od 0 do 1, gdzie dolna granica odpowiada sieciom, w których nie ma relacji, a górna granica przedstawia sieci z wszystkimi możliwymi relacjami. Im bliżej wartości 1, tym bardziej gęsta jest sieć, tym bardziej zwarte są węzły w tej sieci.

Informacje w gęstych sieciach mogą łatwiej przepływać niż informacje w rzadkich sieciach. Tabela „Rzadkie i gęste sieci” przedstawia dwie sieci składające się z siedmiu węzłów. Rzadka sieć zawiera tylko siedem z możliwych 42 relacji między węzłami, dając gęstość o wartości 0,17. Z drugiej strony gęsta sieć zawiera wszystkie możliwe relacje i ma gęstość o wartości 1,0.

Tabela 3. Rzadkie i gęste sieci.



Aby możliwy był przepływ informacji z węzła A do węzła G w rzadkiej sieci, informacja ta musi przejść przez pięć innych węzłów. Z kolei w gęstej sieci informacja może przejść bezpośrednio z węzła A do węzła G.

Stopień węzłowy

Ważne osoby w sieci to często te, które są zaangażowane w największą liczbę relacji. Osoby te uzyskują informacje z różnych źródeł i rozpowszechniają je dużej liczbie innych osób. Z kolei osoby zaangażowane w niewielką liczbę relacji nie mogą bezpośrednio wpływać na dużą liczbę innych osób w sieci.

Stopień węzła, określony jako łączna liczba relacji związanych z tym węzłem, pozwala na porównania między uczestnikami sieci. Osoby, dla których wartość stopnia jest wyższa, są bardziej aktywne, niż osoby z niższą wartością. Stopień ignoruje kierunek relacji, dając ogólną miarę aktywności węzła.

W przypadku ukierunkowanych relacji, podczas zestawiania liczby relacji, możliwe jest skoncentrowanie się na kwestii tego, czy węzeł jest źródłem czy celem. *Stopień wejściowy* dla węzła jest liczbą relacji, w których dany węzeł jest celem. Z kolei *stopień wyjściowy* jest liczbą relacji, w których węzeł jest źródłem. Poniższa tabela przedstawia wartości stopni, stopni wejściowych i wyjściowych dla każdego węzła na rysunku „Przykład sieci społecznościowej”.

Tabela 4. Przykładowe wartości stopnia, stopnia wejściowego i wyjściowego.

Węzeł	Stopień	Stopień wejściowy	Stopień wyjściowy
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



Stopień wejściowy jest często traktowany jako miara prestiżu. Wyższe wartości stopni wejściowych odpowiadają większej liczbie relacji kończących się w tym węźle. Innymi słowy z tymi osobami kontaktuje się duża liczba innych osób. Wiele innych węzłów nawiązuje relacje z tym węzłem. Z kolei stopień wyjściowy jest traktowany jako miara centralności. Wyższe wartości odpowiadają większej liczbie relacji pochodzących z tego węzła. Osoby te kontaktują się z dużą liczbą innych osób.

W przypadku węzłów w przykładowej sieci, wartości stopnia wskazują, że węzły A i D są najaktywniejsze, podczas gdy węzły B i E są najmniej aktywne. Wartości stopni wejściowych wskazują, że węzeł G ma najwyższy prestiż. Na podstawie wartości stopnia wyjściowego można stwierdzić, że węzeł A jest w największym stopniu węzłem centralnym.

Węzły IBM SPSS Modeler Social Network Analysis

Wraz z wieloma standardowymi węzłami, dostarczonymi z IBM SPSS Modeler, użytkownik może pracować także z węzłami IBM SPSS Modeler Social Network Analysis, aby uwzględnić wyniki analizy sieci społecznościowej w swoich strumieniach. Tabela „Węzły IBM SPSS Modeler Social Network Analysis” opisuje te węzły, które są zapisane w palecie Źródło.

Tabela 5. Węzły IBM SPSS Modeler Social Network Analysis.

Węzeł	Ikona	Opis
Analiza grup		Węzeł Analizy grupy importuje dane rekordów dotyczących rozmów telefonicznych z pliku tekstowego ze stałymi polami, identyfikuje grupy węzłów w ramach sieci, zdefiniowane przez te rekordy i generuje kluczowe wskaźniki wydajności dla grup ludzi w tej sieci. Więcej informacji zawiera temat “Przegląd analizy grup” na stronie 9.
Analiza przenikania		Węzeł Analizy przenikania importuje dane rekordów dotyczących rozmów telefonicznych z pliku tekstowego ze stałymi polami, powiela wpływ w ramach sieci, zdefiniowany przez te rekordy i generuje kluczowe wskaźniki wydajności podsumowując wyniki dotyczące wpływu ludzi na poszczególne węzły. Więcej informacji zawiera temat “Przegląd analizy przenikania” na stronie 15.

Zakładki węzłów

Węzły IBM SPSS Modeler Social Network Analysis mają następujące zakładki do definiowania i przeglądania analizy:

- **Zakładka Dane.** Służy do identyfikowania pliku zawierającego informacje o sieci społecznościowej.
- **Zakładka Opcje kompilacji.** Służy do definiowania ustawień dla analiz.
- **Zakładka Analiza.** Służy do przeglądania podsumowania wstępnego wyniku, który zawiera wskazówki dotyczące modyfikacji ustawień danych wejściowych w zakładce Dane w celu zapewnienia optymalnych wyników.

Ponadto węzły zawierają następujące zakładki, wspólne we wszystkich węzłach źródłowych w IBM SPSS Modeler:

- **Zakładka Filtr.** Służy do eliminacji lub zmiany nazw zmiennych wyjściowych stworzonych przez te węzły. Ta zakładka oferuje taką samą funkcjonalność jak węzeł filtrowania.
- **Zakładka Typy.** Służy do ustawiania poziomów pomiaru dla zmiennych wyjściowych stworzonych przez te węzły. Ta zakładka oferuje taką samą funkcjonalność jak węzeł typu.
- **Zakładka Adnotacje.** Służy do zmiany nazw węzłów, zapewnienia definiowanej przez użytkownika podpowiedzi oraz zapisywania długich adnotacji.

Więcej informacji na temat popularnych zakładek znajduje się w dokumentacji IBM SPSS Modeler.

Przeglądanie wyników

Zwykle jako dane wejściowe dla modelu predykcyjnego wykorzystuje się dane wyjściowe z węzłów. Chcąc zobaczyć dane wyjściowe, można dodać do strumienia węzeł Tabela. Jednak ilość danych jest zwykle skrajnie duża, co powoduje, że generowanie tabeli zajmuje dużo czasu. Ponadto duża liczba wierszy w tabeli ogranicza jej użyteczność.

Alternatywnie można wygenerować podgląd tabeli, który wyświetli próbkę danych wyjściowych, które zostaną utworzone. Podgląd pokazuje wygenerowane pola dla ograniczonej liczby wierszy. Liczbę wierszy definiuje się we właściwościach strumienia. Aby zobaczyć podgląd danych wyjściowych węzła, należy wykonać następujące kroki:

1. Otwórz węzeł.

2. Określ ustawienia danych w zakładce Dane.
3. Zdefiniuj parametry analityczne w zakładce Opcje kompilacji.
4. Kliknij opcję **Podgląd**.

Otworzy się okno podglądu, w którym zostaną wyświetlone wyniki. Ponadto przeglądanie danych wyjściowych wypełnia zakładkę Analiza węzła przeglądem podsumowania wyników.

Analizowanie danych

Określenie ustawień analitycznych, które zapewnią optymalne wyniki jest zwykle powtarzającym się procesem. Definiuje się ustawienia, uruchamia analizę i przegląda wyniki. Jeśli wyniki nie są tak bardzo przydatne, jak mogłyby być, należy zmienić ustawienia i ponownie uruchomić analizę.

Aby przeanalizować dane wejściowe węzła, należy wykonać następujące kroki:

1. Otwórz węzeł.
2. Określ ustawienia danych w zakładce Dane.
3. Zdefiniuj parametry analityczne w zakładce Opcje kompilacji.
4. Kliknij opcję **Analizuj dane**.

Jeśli opcje kompilacji wskazują, że powinien się wyświetlać statystyki podsumowujące, zakładka Analiza pokazuje wyniki.

Chcąc ponownie uruchomić analizę, należy kliknąć opcję **Wyczyść analizę**, aby wyczyścić bieżące wyniki przed kliknięciem opcji **Analizuj dane**.

Zastosowania

Konkretne zastosowania, w których oprogramowanie IBM SPSS Modeler Social Network Analysis może być szczególnie przydatne to:

- **Przewidywanie odejść klientów.** Cechy grupowe mogą wpływać na współczynnik odejść. Skoncentrowanie się na osobach w grupach, wykazujące zwiększone ryzyko odejścia, może pozwolić na zapobieganie ich odejściu. Ponadto użytkownik może zidentyfikować osoby, wykazujące ryzyko odejścia dzięki przepływowi informacji od osób, które już odeszły.
- **Wpływanie na liderów grup.** Liderzy grup w znacznym stopniu wpływają na innych członków grupy. Jeśli uda się nie dopuścić do odejścia lidera grupy, może zostać zmniejszony współczynnik odejść dla członków grupy. Alternatywnie próba skłonienia lidera grupy do odejścia od konkurencji może zwiększyć współczynnik odejść członków grupy powiązanych z tą konkurencją, zmniejszając jednocześnie współczynnik odejść członków grupy, powiązanych z Twoją firmą.
- **Marketing.** Możliwe jest wykorzystanie liderów grup do zapoczątkowania propozycji na nowe towary lub usługi. Wpływ lidera może zwiększyć prawdopodobieństwo zakupu proponowanego produktu lub usługi przez innych członków grupy. Można wykorzystać analizę przenikania do zidentyfikowania osób, które są pod największym wpływem liderów grup i skupić na nich działania marketingowe.

Z produktem IBM SPSS Modeler Social Network Analysis dostarczane są dwa strumienie demonstracyjne, które stanowią przykłady sposobu uwzględniania wyników analizy sieci społecznościowych w strumieniach. Pliki danych i przykładowe strumienie są instalowane w folderze *Demos*, w katalogu instalacyjnym produktu.

- *DA demo streams.str* stanowi przykład analizy danych stosowanej w celu zidentyfikowania 300 osób, dla których prawdopodobieństwo odejścia jest najwyższe.
- *GA demo streams.str* stanowi przykład użycia kluczowych wskaźników wydajności (KPI) w celu przewidzenia odchodzenia grup i osób, a także użycia wskaźników KPI w celu doboru osób do grupy docelowej dla kampanii marketingowej.

Struktura danych

Informacje o interesujących osobach w sieci mogą być znajdować się w różnych plikach, bazach danych i systemach w firmie użytkownika. Aby analizować sieć przy pomocy oprogramowania IBM SPSS Modeler Social Network Analysis, trzeba wydobyc odpowiednie rekordy i pola ze źródeł danych wejściowych i sformatować je tak, aby można je było zaimportować do węzłów.

Węzły analityczne wymagają rekordów dotyczących rozmów telefonicznych, zapisanych w pojedynczym pliku tekstowym, o stałej szerokości tekstu. Każdy wiersz pliku odpowiada relacji, gdzie dane są zorganizowane w następujących kolumnach:

- identyfikator osoby, która rozpoczyna relację.
- identyfikator osoby, która jest celem relacji.
- opcjonalna waga relacji.

Wszystkie dane muszą być numeryczne, a identyfikatory osób muszą być liczbami całkowitymi. Istnieje możliwość dodania w pliku nazw pól w pierwszym wierszu danych. Tabela „Przykład rekordów danych połączeń” ilustruje tę strukturę danych.

W rekordach nagłówkowych także muszą być używane albo liczby, albo wartości puste. Na przykład nagłówek z wartością *EF BB BF* spowoduje błąd.

Tabela 6. Przykład rekordów danych dotyczących rozmów telefonicznych.

Źródło	Cel	Waga
1000	5642	243
2190	8444	831
0299	9419	559

Wartości wagi mogą odpowiadać dowolnej mierze, która ma być wykorzystywana do przedstawiania znaczenia relacji względem innych relacji w sieci. W przypadku danych dotyczących rozmów telefonicznych, popularne wagi obejmują czas trwania rozmowy lub częstotliwość rozmów. Należy zwrócić uwagę na to, że obowiązuje to w przypadku analizy dyfuzji, ale analiza grupy obsługuje tylko częstotliwości rozmów.

Jeśli użytkownik chce, aby analiza koncentrowała się na podzbiorze historii rozmów telefonicznych, musi on użyć tego podzbioru podczas tworzenia pliku z danymi wejściowymi. Na przykład można ograniczyć analizę do ostatnich kilku miesięcy lub do ostatnich rozmów telefonicznych z daną osobą, włączając tylko te dane do pliku tekstowego z danymi wejściowymi.

Właściwości skryptów

Skrypty w IBM SPSS Modeler Social Network Analysis są potężnym narzędziem automatyzacji procesów interfejsu użytkownika. Skrypty mogą wykonywać takie same działania, które wykonujesz ty używając myszy lub klawiatury, możesz je wykorzystać do zautomatyzowania wysoce powtarzalnych zadań lub takich zadań, których wykonanie ręczne wymaga dużo czasu. Wyjaśnienie, w jaki sposób używać skryptów można znaleźć w podręczniku *ScriptingAutomation.pdf* dostępnym razem z IBM SPSS Modeler.

Właściwości węzła przenikania

W poniższej tabeli znajdują się właściwości skryptów dla Węzła przenikania.

Nazwa właściwości	Typ danych	Opis właściwości
input_data_file_name	<i>string</i>	
calling_field	<i>field</i>	

Nazwa właściwości	Typ danych	Opis właściwości
called_field	<i>field</i>	
frequency_weight_field	<i>field</i>	
read_field_names	<i>boolean</i>	
diffusion_list_file_name	<i>string</i>	
spreading_factor	<i>double</i>	Wartość domyślna = 50. Min = 1 Maks = 99.
max_number_iterations	<i>integer</i>	Wartość domyślna = 100. Min = 1.
accuracy_threshold	<i>double</i>	Wartość domyślna = 0,01. Min = 0,001.
calculate_statistics	<i>boolean</i>	

Właściwości węzła Analizy grup

W poniższej tabeli znajdują się właściwości skryptów dla Węzła analizy grup.

Nazwa właściwości	Typ danych	Opis właściwości
input_data_file_name	<i>string</i>	
calling_field	<i>field</i>	
called_field	<i>field</i>	
frequency_weight_field	<i>field</i>	
read_field_names	<i>boolean</i>	
coverage_threshold	<i>double</i>	Wartość domyślna = 10. Min = 1 Maks = 99.
min_group_size	<i>integer</i>	Wartość domyślna = 2. Min = 2.
max_group_size	<i>integer</i>	Wartość domyślna = 100. Min = 2.
calculate_statistics	<i>boolean</i>	

Rozdział 2. Analiza grup

Przegląd analizy grup

Analiza grup wykorzystuje wzorce interakcji pojedynczych osób w sieci w celu identyfikacji grup podobnych osób. Charakterystyka tych grup wpływa na zachowanie poszczególnych członków grup. Na przykład niewielkie grupy posiadające wiele relacji między członkami oraz silnych liderów charakteryzują się zwiększonym ryzykiem odejść, nawet jeśli żadne z członków grupy jeszcze nie odszedł.³ Modele predykcyjne obejmujące grupowe i indywidualne mierniki dadzą lepsze wyniki, niż modele obejmujące tylko te drugie mierniki.

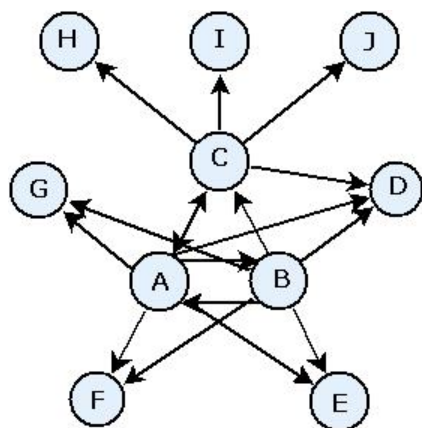
Analiza grup składa się z następujących, ogólnych kroków:

1. Określenie siły relacji, która odzwierciedla sąsiedztwo społeczne. Więcej informacji zawiera temat “Określenie podobieństw społecznych”.
2. Podzielenie sieci na grupy w oparciu o siłę relacji, z uwzględnieniem ograniczeń rozmiaru. Więcej informacji zawiera temat “Dzielenie na grupy” na stronie 10.
3. Stworzenie profili grup i pojedynczych osób, z uwzględnieniem identyfikacji liderów grup. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.

Określenie podobieństw społecznych

Członkowie grupy powinni być bardziej podobni do siebie nawzajem, niż do osób, które nie należą do grupy. W analizie sieci podobieństwo dwóch węzłów zależy od ich relacji. Dla każdego węzła w sieci występuje zestaw węzłów, które są celami ukierunkowanych relacji z węzłem. W przypadku danych telekomunikacyjnych zestaw ten odpowiada wszystkim ludziom, z którymi kontaktowała się określona osoba. Jeśli dwie osoby kontaktują się z tym samym zestawem ludzi, osoby te są uznawane za podobne do siebie. Im więcej zestawów celów relacji nakłada się na siebie dla dwóch osób, tym bardziej są one do siebie podobne.

Weźmy pod uwagę sieć pokazaną na rysunku „Przykład sieci składającej się z dziesięciu węzłów”.



Rysunek 2. Przykład sieci składającej się z dziesięciu węzłów

Tabela „Węzły docelowe” przedstawia węzły docelowe dla relacji rozpoczynających się w węzłach A, B i C.

3. Richter, Y., E. Yom-Tov i N. Slonim. 2010. Predicting customer churn in mobile networks through analysis of social groups. W: *Proceedings of the 2010 SIAM international conference on data mining*. Columbus, OH: SDM 2010.

Tabela 7. Węzły docelowe.

Węzeł źródłowy	Węzły docelowe
A	B, C, D, E, F, G
B	A, C, D, E, F, G
C	A, D, H, I, J

Węzły A i B mają pięć wspólnych węzłów docelowych. Z drugiej strony węzły A i C mają tylko jeden. W ten sposób węzły A i B są bardziej podobne społecznie, niż węzły A i C.

Wagi relacji, takie jak czas trwania lub częstotliwość rozmów telefonicznych nie obejmują podobieństwa węzłów w sieci. Aby uwzględnić pojęcie podobieństwa, analizy grup wykorzystują wspólne informacje⁴ jako wagę relacji. Ten parametr odzwierciedla prawdopodobieństwo powiązania dwóch węzłów z tym samym, innym węzłem. Jeśli relacje w sieci mają wstępnie zdefiniowane wagi, takie jak częstotliwość rozmów telefonicznych w przypadku danych telekomunikacyjnych, wspólne informacje obejmują odpowiednio te wagi.

Dzielenie na grupy

Grupy powinny się składać z osób o wysokiej wadze relacji z innymi członkami grupy, gdzie wagi mierzą podobieństwo węzłów w relacji. W wyniku tego identyfikacja grupy rozpoczyna się od pominięcia słabszych relacji w sieci. *Wartość progowa pokrycia* kontroluje ten proces poprzez definiowanie udziału utrzymujących się, najsilniejszych relacji. Na przykład wartość progowa 0,4 dla pokrycia skutkuje tym, że do identyfikacji grupy wykorzystywanych jest 40% najsilniejszych relacji, a pozostałe 60% relacji jest pomijanych.

Pozostałe relacje mogą dać pewne bardzo małe lub bardzo duże grupy, które w ograniczonym stopniu nadają się do przewidywania. Aby zapobiegać uwzględnianiu takich grup w analizie można określić minimalne i maksymalne rozmiary grup. Grupy o rozmiarach mniejszych od minimalnego rozmiaru są całkowicie pomijane. Jednak grupy o rozmiarach większych od maksymalnego są dzielone na mniejsze grupy z zakresu rozmiarów, które są do przyjęcia. Grupy, które zostaną po odrzuceniu słabych relacji i narzuceniu ograniczeń co do rozmiaru, są nazywane *grupami kluczowymi*.

Usunięcie relacji z oryginalnej sieci może skutkować tym, że niektóre osoby nie znajdują się w żadnej kluczowej grupie. Jednak osoby te mogą mieć powiązania z członkami grupy, które będą gwarantowały włączenie do grupy. Osoba zostaje dodana do grupy, jeśli istnieje wiele względnie silnych relacji z kluczowymi członkami tej grupy, przy założeniu, że nie zostanie naruszone ograniczenie względem rozmiaru grupy. W ten sposób ostateczne grupy składają się z zestawu kluczowych członków oraz z członków dodanych ze względu na powiązania z rdzeniem.

Opisywanie grup i członków grup

Poza gęstością, stopniami wejściowymi i stopniami wyjściowymi, inne parametry opisują dynamikę grupy. W szczególności wynik autorytetu i rozpowszechniania daje możliwość zmierzenia statusu społecznego poszczególnych osób w grupach. Rola każdej osoby w grupie jest szczególnie ważna w przypadku próby przewidywania zachowania grupy oraz jej członków.

Wynik autorytetu dla węzła mierzy tendencję innych węzłów w grupie do łączenia się z nim. Jeśli wiele osób kontaktuje się z konkretną osobą, potencjalnie prosząc o informacje lub opinie, osoba ta ma rolę autorytetu. Wyniki autorytetu dla węzłów w grupie, które odpowiadają próbom łączenia się z losowo wybranymi węzłami (za każdym razem są to nowe węzły) i sprawdzania, czy węzły te należą do grupy, mieszczą się w przedziale od 0 do 1. Im wynik autorytetu jest bliższy wartości 1, tym większym autorytetem cieszy się ten węzeł w grupie. Węzeł w grupie z najwyższym wynikiem autorytetu nazywany jest *liderem w zakresie autorytetu* dla grupy. Po podzieleniu najwyższego wyniku w grupie przez najniższy otrzymuje się miarę łącznej siły lidera w zakresie autorytetu.

4. Cover, T. M. i J. A. Thomas. 2006. *Elements of Information Theory*, wyd. 2. New York: John Wiley and Sons, Inc.

Z kolei *wynik rozpowszechniania* dla węzła mierzy tendencję węzła do łączenia się z innymi węzłami w grupie. Dana osoba kontaktuje się z wieloma innymi osobami w grupie, osoba ta może w znacznym stopniu wpływać na opinie całej grupy. Wyniki rozpowszechniania dla węzłów w grupie, które odpowiadają próbom łączenia się z losowo wybranymi węzłami (za każdym razem są to nowe węzły) i sprawdzania, czy węzły te należą do grupy, mieszczą się w przedziale od 0 do 1. Im wynik rozpowszechniania jest bliższy wartości 1, tym bardziej węzeł ten łączy się z innymi członkami grupy. Węzeł w grupie z najwyższym wynikiem rozpowszechniania nazywany jest *liderem rozpowszechniania* dla grupy. Po podzieleniu najwyższego wyniku w grupie przez najniższy otrzymuje się miarę łącznej siły lidera rozpowszechniania.

Węzeł Analizy grup

Węzeł Analizy grup, który jest dostępny z palety Źródła, identyfikuje zestawy osób w sieci, które są społecznie podobne do siebie i określa pokrewny status społeczny poszczególnych osób w grupach. Grupy, które są względnie niewielkie, mają wiele powiązań między ich członkami i posiadają silnego lidera, cechują się zwiększonym ryzykiem odejść z uwagi na dynamikę grupy. Można wykorzystać te wyniki do określenia wyników odejścia dla grupy, które w połączeniu z wynikami odejść dla poszczególnych osób, polepszają możliwość przewidzenia odejść w porównaniu z modelami opartymi wyłącznie na wynikach poszczególnych osób.

Jeśli się uwzględni ilość danych, z których normalnie składa się sieć i naturę interakcji społecznych podlegających definicjom grupy, analiza grup jest procesem czasochłonnym. Zwykle wyniki analizy grup zapisywane są w bazie danych lub pliku, który ma zostać użyty jako dane wejściowe do modelu predykcyjnego. Wyniki grupy były by odświeżane stosunkowo rzadko, mniej więcej raz w miesiącu, podczas gdy model może być odświeżany znacznie częściej.

Wymagania. Węzeł wymaga pliku tekstowego o stałej szerokości tekstu, definiującego sieć społecznościową za pomocą trzech pól. Jedno pole identyfikuje źródło dla każdej relacji ukierunkowanej, jedno pole definiuje cel dla każdej relacji ukierunkowanej, a trzeci plik określa opcjonalną siłę dla każdej relacji. Wszystkie relacje w sieci muszą być ukierunkowane.

Określenie danych do analizy grup

Zakładka Dane okna węzła źródłowego Analiza grup umożliwia określenie pliku wejściowego, zawierającego relacje węzła sieciowego.

Plik. Należy określić nazwę pliku lub folderu zawierającego rekordy dotyczące rozmów telefonicznych. Można wpisać nazwę lub kliknąć przycisk elipsy (...), aby wybrać nazwę z systemu plików. Ścieżka zostanie pokazana po wybraniu nazwy, a jej zawartość zostanie wyświetlona z ogranicznikami. Jeśli określisz folder, rekordy dotyczące rozmów telefonicznych we wszystkich plikach tego folderu zostają połączone do celów analizy; wszystkie pliki w folderze powinny mieć tę samą strukturę.

Wczytaj nazwy pól z pliku. Opcja ta, która jest wybrana domyślnie, traktuje pierwszy wiersz w pliku z danymi jako nazwy kolumn. Jeśli pierwszy wiersz nie jest nagłówkiem, należy odznaczyć tę opcję, aby automatycznie nadać każdemu polu nazwę ogólną, taką jak *Pole1* i *Pole2*.

Ustawienie definicji sieci

Ustawienia definicji sieci definiują role pól.

Pola. Aby ręcznie przypisać pozycje z tej listy do różnych pól ról, należy użyć klawiszy strzałek. Ikony wskazują prawidłowe poziomy pomiaru dla każdego pola roli. Kliknij przycisk **Wszystkie**, aby wybrać wszystkie pola na liście, lub kliknij przycisk poziomu pojedynczego pomiaru, aby wybrać wszystkie pola, które posiadają ten poziom pomiaru.

Źródło. Wybierz jedno pole jako źródło ukierunkowanej relacji.

Cel. Wybierz jedno pole jako cel ukierunkowanej relacji.

Waga. Opcjonalnie można wybrać jedno pole odpowiadające liczbie wystąpień relacji w sieci. Analiza odpowiednio waży rekordy podczas określania siły relacji.

Ustawianie opcji kompilacji dla analizy grup

Zakładka Opcje kompilacji okna węzła źródłowego Analiza grup pozwala na definiowanie opcji do identyfikowania grup w sieci.

Ustawienia analizy grup

Ustawienia analizy grup mają wpływ na rozmiar i względną siłę tworzonych grup.

Wartość progowa pokrycia. Należy zdefiniować udział najsilniejszych relacje w sieci, które mają zostać użyte w analizie. Na przykład wartość progowa pokrycia o wartości 0,2 wskazuje, że do analizy zostaną wykorzystane tylko te relacje, których waga należy do 20% najwyższych wartości. Parametr ten, mieszczący się w przedziale od 0 do 1, pozwala na to, aby analiza koncentrowała się na najsilniejszych relacjach, które powinny występować w grupach. Wyższe wartości progowe dają grupy, w których relacje między członkami grupy są słabsze. Więcej informacji zawiera temat “Dzielenie na grupy” na stronie 10.

Minimalny rozmiar grupy. Należy podać dolną granicę rozmiaru grupy. Grupy mniejsze od tej wartości nie zostaną zwrócone w wynikach.

Maksymalny rozmiar grupy. Należy podać górną granicę rozmiaru grupy. Grupy większe od tej wartości zostaną podzielone na mniejsze grupy.

Oblicz i wyświetlaj statystyki podsumowujące. Jeśli wybierze się tę opcję, poza uzyskaniem danych wyjściowych dla kluczowych wskaźników wydajności, węzeł obliczy i wyświetli statystyki podsumowujące dla analizy. Obliczenie tego parametru może negatywnie wpłynąć na wydajność węzła w przypadku skrajnie dużych sieci.

Przeglądanie statystyki analizy grup

Zakładka Analiza okna węzła źródłowego Analiza grup udostępnia przegląd podsumowania grup zidentyfikowanych w sieci. Tabela „Statystyka podsumowująca Analiza grup” zawiera dostępne statystyki podsumowujące dla grup.

Tabela 8. Statystyka podsumowująca Analiza grup.

Statystyka	Opis
Łączna liczba węzłów w grupach	Liczba węzłów zawartych w zidentyfikowanych grupach
Całkowita liczba łączy w grupach	Liczba łączy zawartych w zidentyfikowanych grupach
Łączna liczba grup	Liczba grup zidentyfikowanych w sieci
Średni rozmiar grupy	Średnia liczba węzłów w grupie
Średnia gęstość grupy	Średni udział bezpośrednich powiązań między węzłami w grupie. Więcej informacji zawiera temat “Gęstość sieci” na stronie 3.
Średnia ilość energii kluczowego członka	Średni udział węzłów w grupie, które są węzłami kluczowymi dla grupy. Więcej informacji zawiera temat “Dzielenie na grupy” na stronie 10.
Średnia gęstość grupy kluczowej	Średni udział bezpośrednich powiązań między kluczowymi węzłami w grupie.
Średni stopień wejściowy	Średnia liczba połączeń przychodzących. Więcej informacji zawiera temat “Stopień węzłowy” na stronie 4.
Średni stopień wyjściowy	Średnia liczba połączeń wychodzących. Więcej informacji zawiera temat “Stopień węzłowy” na stronie 4.

Należy wybrać określony parametr z tabeli Statystyki podsumowujące, aby zobaczyć rozkład wartości, odchylenie standardowe oraz współczynnik skośności parametru.

Sprawdzenie tych parametrów pomaga w definiowaniu ustawień analizy grup. W idealnej sytuacji wyniki powinny pokazywać, że stosunkowo niewielkie grupy mają wysokie wartości gęstości. Na przykład jeśli istnieją jakieś bardzo duże grupy, należy rozważyć zmniejszenie maksymalnego rozmiaru grupy, aby podzielić je na mniejsze grupy. Alternatywnie jeśli wartości gęstości są niskie, należy rozważyć zmniejszenie wartości progowej pokrycia, aby skupić się na silniejszych relacjach.

Dane wyjściowe do analizy grup

Węzeł Analizy grup generuje różne pola opisujące grupy i osoby w tych grupach. Możliwe jest powiększenie istniejących modeli i danych o te kluczowe wskaźniki wydajności w celu polepszenia predykcji generowanych przez modele użytkownika. Na przykład istnieje możliwość aktualizacji wartości skłonności pojedynczych osób do odejść w celu uwzględnienia wpływu grup.

Tabela „Kluczowe wskaźniki wydajności dla grup” zawiera pola z kluczowymi wskaźnikami wydajności dla grup zidentyfikowanych w analizie.

Tabela 9. Kluczowe wskaźniki wydajności dla grup.

Pole	Opis
GAG_GroupNumber	Unikalny identyfikator grupy
GAG_Size	Liczba osób w grupie
GAG_Density	Udział bezpośrednich powiązań między osobami w grupie. Więcej informacji zawiera temat “Gęstość sieci” na stronie 3.
GAG_KernelDensity	Udział bezpośrednich powiązań między kluczowymi osobami w grupie
GAG_CoreNodesFraction	Udział osób w grupie, które są osobami kluczowymi dla grupy. Więcej informacji zawiera temat “Dzielenie na grupy” na stronie 10.
GAG_MaxRankType1	Maksymalny wynik autorytetu dowolnego członka grupy. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAG_MinRankType1	Minimalny wynik autorytetu dowolnego członka grupy
GAG_MaxMinRankRatioType1	Współczynnik najwyższego wyniku autorytetu do najniższego. Wartość ta odzwierciedla siłę autorytetu lidera grupy.
GAG_MaxRankType2	Maksymalny wynik rozpowszechnienia dowolnego członka grupy. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAG_MinRankType2	Minimalny wynik rozpowszechnienia dowolnego członka grupy
GAG_MaxMinRankRatioType2	Współczynnik najwyższego wyniku rozpowszechnienia do najniższego. Wartość ta odzwierciedla siłę rozpowszechnienia lidera grupy.

Tabela „Kluczowe wskaźniki wydajności dla osób” zawiera pola z kluczowymi wskaźnikami wydajności dla osób w sieci.

Tabela 10. Kluczowe wskaźniki wydajności dla osób.

Pole	Opis
GAI_NodeNumber	Unikalny identyfikator osoby
GAI_CoreNode	Wskaźnik określający, czy dana osoba jest osobą kluczową dla grupy, czy nie. Więcej informacji zawiera temat “Dzielenie na grupy” na stronie 10.
GAI_RankType1	Wynik autorytetu dla danej osoby. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAI_RankOrderType1	Kolejność rankingu w grupie w oparciu o wyniki autorytetu
GAI_RankType2	Wynik rozpowszechnienia dla danej osoby. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAI_RankOrderType2	Kolejność rankingu w grupie w oparciu o wyniki rozpowszechnienia

Tabela 10. Kluczowe wskaźniki wydajności dla osób (kontynuacja).

Pole	Opis
GAI_InDegree	Liczba relacji, w których dana osoba jest celem relacji. Więcej informacji zawiera temat “Stopień węzłowy” na stronie 4.
GAI_OutDegree	Liczba relacji, w których dana osoba jest źródłem relacji. Więcej informacji zawiera temat “Stopień węzłowy” na stronie 4.
GAI_GroupLeaderType1	Czy węzeł jest liderem w zakresie autorytetu, którego wynik przywództwa jest pochodną przychodzących połączeń. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAI_GroupLeaderConfidenceType1	Pewność, że węzeł jest liderem w zakresie autorytetu.
GAI_GroupLeaderType2	Czy węzeł jest liderem rozpowszechniania, którego wynik przywództwa jest pochodną wychodzących połączeń. Więcej informacji zawiera temat “Opisywanie grup i członków grup” na stronie 10.
GAI_GroupLeaderConfidenceType2	Pewność, że ten węzeł jest liderem w zakresie rozpowszechniania.

Rozdział 3. Analiza przenikania

Przegląd analizy przenikania

Analiza przenikania identyfikuje osoby, które są najbardziej podatne na wpływ innych osób w sieci społecznościowej, mierząc wpływ jako *rozproszoną energię*. W procesie stosowana jest strategia rozprzestrzeniającej się aktywacji, w której wpływ iteracyjnie rozprzestrzenia się z węzłów sieci do węzłów bezpośrednio z nimi sąsiadujących, i zmniejsza się, przechodząc od węzła do węzła.⁵ Po otrzymaniu energii węzeł staje się aktywny i przesyła część tej energii do wszystkich, sąsiednich węzłów, będących celem ukierunkowanych relacji z tym węzłem.

Współczynnik rozprzestrzeniania definiuje stosunek energii przesyłanej przez aktywny węzeł, do energii, która pozostaje w tym węźle. Wszystkie węzły, które otrzymują tę energię same prześlą tę samą ilość energii sąsiadującym węzłom, co będzie skutkowało procesem rozpadu przesyłanej energii. Duże współczynniki rozprzestrzeniania oznaczają przesyłanie większej ilości energii, co pozwala na dotarcie energii do węzłów, które są bardziej odległe od pierwotnie aktywowanych węzłów przed całkowitym rozpadem tego procesu. Niskie współczynniki rozprzestrzeniania skutkują szybkim rozpadem procesów przenikania, gdzie przesyłana energia pozostaje stosunkowo blisko początkowych węzłów.

Łączna ilość energii rozproszonej przez aktywowany węzeł jest rozprowadzana wśród wszystkich węzłów będących celem ukierunkowanych relacji z tym węzłem. Ilość energii, którą otrzymuje każdy węzeł zależy od siły relacji z aktywowanym węzłem. Ilość przekazana do konkretnego węzła jest równa wadze relacji podzielonej przez sumę wag wszystkich relacji, dla których dany aktywowany węzeł jest źródłem. Dlatego węzły sąsiadujące, które mają relacje o wyższych wagach względnych, otrzymują większą ilość energii niż węzły sąsiadujące z niższymi wagami względnymi.

Proces przenikania zatrzymuje się, gdy wystąpi jeden z poniższych warunków:

- aktywowane węzły nie są źródłem żadnych ukierunkowanych relacji
- ilość przesyłanej energii znajduje się poniżej *wartości progowej dokładności*, granicy ilości przesyłanej energii, która wystarcza do kontynuowania procesu
- liczba powtórzeń osiąga określony limit

Gdy przenikanie kończy się, węzły posiadające najwięcej rozproszonej energii są najbardziej wrażliwe na wpływ, który rozpoczął ten proces. Na przykład jeśli proces rozpoczyna się od węzłów, które odchodzą, węzły o największej ilości energii same są najbardziej narażone na ryzyko odejścia. Można poświęcić tym węzłom szczególną uwagę, aby nie dopuścić do ich odejścia.

Przykład analizy przenikania

Weźmy pod uwagę sieć pokazaną na "Pokazywanie sieci" na stronie 2. Sieć ta składa się z siedmiu węzłów posiadających wzajemne, ukierunkowane relacje o różnej sile.

Tabela „Przykład procesu przenikania” przedstawia proces przenikania w sieci, gdzie współczynnik rozprzestrzeniania wynosi 0,80. Początkowo węzeł A zawiera całą energię, arbitralnie przypisaną wartość 1,00. W kroku 1, węzeł ten aktywuje, rozprzestrzeniając 80% swojej energii do trzech sąsiadujących węzłów, które są celami relacji i zachowuje 20% energii dla siebie. Relacje te mają łączną wagę o wartości 100. Relacja z węzłem D zaliczana jest jako połowa tej łącznej wartości tak, że węzeł ten otrzymuje połowę rozproszonej energii, lub 0,40. Relacja z węzłem B zaliczana jest jako 20% łącznej wagi, dając rozproszoną energię o wartości 0,16. Węzeł C otrzymuje pozostałą ilość energii, 0,24, która stanowi 30% energii pochodzącej z węzła A.

5. Dasgupta, K., R. Singh, B. Viswanathan, D. Chakraborty, S. Mukherjee, A. A. Nanavati i A. Joshi. 2008. Social ties and their relevance to churn in mobile telecom networks. W: *Proceedings of the 11th international conference on extending database technology*. New York, NY: ADM.

Tabela 11. Przykład procesu przenikania.

Krok	Węzeł A	Węzeł B	Węzeł C	Węzeł D	Węzeł E	Węzeł F	Węzeł 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

Po otrzymaniu energii w kroku 1, węzły B, C i D aktywują się w kroku 2 i rozpraszają energię do wszystkich sąsiadujących węzłów. Węzeł B nie ma węzłów docelowych, zatem nie rozprasza energii. Z drugiej strony węzeł C ma cel i rozprasza 80% swojej energii do tego węzła, G. W końcu węzeł D rozprzestrzenia 80% swojej energii do węzłów E i F, przy czym F otrzymuje trzy razy więcej energii, niż węzeł E z uwagi na silniejszą relację z węzłem D.

W kroku 3 aktywują się węzły E, F i G. Węzły E i G nie mają węzłów docelowych, zatem nie rozpraszają energii. Jednak węzeł F rozprasza 80% swojej energii na rzecz swojego jedyne go węzła docelowego, G. Teraz węzeł ten ma łączną energię o wartości 0,38, na którą składa się energia otrzymana od węzła C i F.

W tym momencie aktywuje się węzeł G, ale nie ma węzłów docelowych, do których mógłby przekazać energię, zatem proces przenikania kończy się. Energia pochodząca od węzła A rozproszyła się w sieci, gdzie węzeł G otrzymał największą ilość energii. Jeśli energia wprowadzona do sieci reprezentuje odchodzącego klienta, węzeł G znalazłby się pod największym wpływem węzła A, który byłby tym odchodzącym klientem.

Węzeł Analizy przenikania

Węzeł Analizy przenikania, dostępny z palety Źródła, powiela wpływ wywierany przez określony zestaw ludzi w sieci społecznościowej, wykorzystując relacje sieciowe do identyfikacji osób, które znalazły się pod największym wpływem. Jeśli wpływ oznacza na przykład zamiar odejścia klienta, węzeł identyfikuje te osoby, które najprawdopodobniej ojadą, ponieważ inne określone osoby w sieci odeszły. Możliwe jest powiększenie istniejących modeli i danych o te dane wyjściowe węzła w celu polepszenia predykcji generowanych przez te modele. Na przykład istnieje możliwość aktualizacji wartości skłonności pojedynczych osób do odejść w celu uwzględnienia wpływu przenikania.

Wymagania. Węzeł wymaga dwóch plików tekstowych o stałej szerokości tekstu, zawierających dane do analizy. Pierwszy plik definiuje sieć społecznościową wykorzystującą trzy pola. Jedno pole identyfikuje źródło dla każdej relacji ukierunkowanej, jedno pole definiuje cel dla każdej relacji ukierunkowanej, a trzeci plik określa opcjonalną siłę dla każdej relacji. Wszystkie relacje w sieci muszą być ukierunkowane. Drugi plik zawiera listę identyfikatorów, od których powinien rozpocząć się wpływ.

Określenie danych do analizy przenikania

Zakładka Dane okna węzła źródłowego Analiza przenikania umożliwia określenie pliku z danymi wejściowymi, zawierającego relacje węzła sieciowego.

Plik. Należy określić nazwę pliku lub folderu zawierającego rekordy dotyczące rozmów telefonicznych. Można wpisać nazwę lub kliknąć przycisk elipsy (...), aby wybrać nazwę z systemu plików. Ścieżka zostanie pokazana po wybraniu nazwy, a jej zawartość zostanie wyświetlona z ogranicznikami. Jeśli określisz folder, rekordy dotyczące rozmów telefonicznych we wszystkich plikach tego folderu zostają połączone do celów analizy; wszystkie pliki w folderze powinny mieć tę samą strukturę.

Wczytaj nazwy pól z pliku. Opcja ta, która jest wybrana domyślnie, traktuje pierwszy wiersz w pliku z danymi jako nazwy kolumn. Jeśli pierwszy wiersz nie jest nagłówkiem, należy odznaczyć tę opcję, aby automatycznie nadać każdemu polu nazwę ogólną, taką jak *Pole1* i *Pole2*.

Ustawienie definicji sieci

Ustawienia definicji sieci definiują role pól.

Pola. Aby ręcznie przypisać pozycje z tej listy do różnych pól ról, należy użyć klawiszy strzałek. Ikony wskazują prawidłowe poziomy pomiaru dla każdego pola roli. Kliknij przycisk **Wszystkie**, aby wybrać wszystkie pola na liście, lub kliknij przycisk poziomu pojedynczego pomiaru, aby wybrać wszystkie pola, które posiadają ten poziom pomiaru.

Źródło. Wybierz jedno pole jako źródło ukierunkowanej relacji.

Cel. Wybierz jedno pole jako cel ukierunkowanej relacji.

Waga. Należy wybrać jedno pole reprezentujące wagę relacji. W przypadku rekordów dotyczących rozmów telefonicznych, waga może odpowiadać jednej z wielu charakterystyk rozmów telefonicznych, która odzwierciedla siłę relacji, taką jak czas trwania rozmowy lub częstotliwość rozmów.

Ustawianie opcji kompilacji dla analizy przenikania

Zakładka Opcje kompilacji okna węzła źródłowego Analiza przenikania pozwala na definiowanie opcji dla procesu przenikania.

Lista ziaren przenikania

Lista ziaren przenikania identyfikuje węzły sieciowe, z których powinno rozpocząć się przenikanie energii. Na przykład w analizie odchodzących klientów lista ta identyfikuje węzły sieciowe, które już odeszły. Alternatywnie lista ta mogłaby zawierać węzły wykazujące najwyższe ryzyko odejścia w celu badania wpływu na sieć w przypadku rzeczywistego odejścia tych węzłów.

Plik. Należy podać nazwę pliku tekstowego zawierającego identyfikatory węzłów, gdzie każdy identyfikator odpowiada jednej linii w pliku. Można wpisać nazwę pliku lub kliknąć przycisk elipsy (...), aby wybrać plik.

Ustawienia analizy przenikania

Ustawienia analizy przenikania określają agresywność przenikania i moment zatrzymania tego procesu.

Współczynnik rozprzestrzeniania. Należy ustalić ilość energii, którą węzeł przesyła w każdym kroku procesu przenikania. Niższe wartości ograniczają energię do najbliższych węzłów w stosunku do węzła, który rozprasza energię. Z kolei wyższe wartości pozwalają, aby energia wpływała na bardziej odległe węzły.

Maksymalna liczba powtórzeń. Określenie górnej granicy dla liczby powtórzeń przenikania. Proces przenikania kończy się automatycznie, gdy ta granica zostanie osiągnięta.

Wartość progowa dokładności. Należy podać najmniejszą zmianę energii węzła, która gwarantuje kontynuowanie procesu przenikania. Jeśli zmiana energii dla każdego węzła jest mniejsza od tej wartości, proces przenikania zakończy się.

Oblicz i wyświetlaj statystyki podsumowujące. Jeśli wybierze się tę opcję, poza uzyskaniem danych wyjściowych dla kluczowych wskaźników wydajności, węzeł obliczy i wyświetli statystyki podsumowujące dla analizy. Obliczenie tego parametru może negatywnie wpłynąć na wydajność węzła w przypadku skrajnie dużych sieci.

Przeglądanie statystyki analizy przenikania

Zakładka Analiza okna węzła źródłowego Analiza przenikania zapewnia przegląd podsumowania wyników przenikania. Tabela „Statystyki podsumowujące Analiza przenikania” zawiera dostępne statystyki podsumowujące.

Tabela 12. Statystyki podsumowujące Analiza przenikania.

Statystyka	Opis
Łączna liczba węzłów w sieci	Liczba węzłów w sieci
Całkowita liczba łączy w sieci	Liczba łączy w sieci
Łączna liczba ziaren przenikania w sieci	Liczba węzłów wykorzystywanych jako ziarna w procesie przenikania

Tabela 12. Statystyki podsumowujące Analiza przenikania (kontynuacja).

Statystyka	Opis
Średni wpływ	Średnia ilość rozproszonej energii powiązanej z danymi osobami.
Średni stopień wejściowy	Średnia liczba relacji, w których dana osoba jest celem relacji. Więcej informacji zawiera temat "Stopień węzłowy" na stronie 4.
Średni stopień wyjściowy	Średnia liczba relacji, w których dana osoba jest źródłem relacji. Więcej informacji zawiera temat "Stopień węzłowy" na stronie 4.

Należy wybrać określony parametr z tabeli Statystyki podsumowujące, aby zobaczyć rozkład wartości, odchylenie standardowe oraz współczynnik skośności parametru.

Dane wyjściowe dla analizy przenikania

Węzeł Analizy przenikania generuje różne pola opisujące osoby w sieci. Możliwe jest powiększenie istniejących modeli i danych o te kluczowe wskaźniki wydajności w celu polepszenia predykcji generowanych przez modele użytkownika. Na przykład istnieje możliwość aktualizacji wartości skłonności pojedynczych osób do odejść w celu uwzględnienia wpływu przenikania.

Tabela „Kluczowe wskaźniki wydajności dla Analizy przenikania” zawiera pola z kluczowymi wskaźnikami wydajności dla osób w analizie.

Tabela 13. Kluczowe wskaźniki wydajności dla Analizy przenikania.

Pole	Opis
DA_NodeNumber	Unikalny identyfikator osoby
DA_DiffusedEnergy	Ilość rozproszonej energii powiązanej z daną osobą. W przypadku analizy odejść, wyższe wartości wskazują większą skłonność do odejść, niż niższe wartości.
DA_InDegree	Liczba relacji, w których dana osoba jest celem relacji. Więcej informacji zawiera temat "Stopień węzłowy" na stronie 4.
DA_OutDegree	Liczba relacji, w których dana osoba jest źródłem relacji. Więcej informacji zawiera temat "Stopień węzłowy" na stronie 4.

Uwagi

Niniejsza publikacja została przygotowana z myślą o produktach i usługach oferowanych w Stanach Zjednoczonych. Materiał ten jest również dostępny w IBM w innych językach. Jednakże w celu uzyskania dostępu do takiego materiału istnieje konieczność posiadania egzemplarza produktu w takim języku.

Produktów, usług lub opcji opisywanych w tym dokumencie IBM nie musi oferować we wszystkich krajach. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela IBM. Odwołanie do produktu, programu lub usługi IBM nie oznacza, że można użyć wyłącznie tego produktu, programu lub usługi IBM. Zamiast nich można zastosować ich odpowiednik funkcjonalny pod warunkiem, że nie narusza to praw własności intelektualnej IBM. Jednakże cała odpowiedzialność za ocenę przydatności i sprawdzenie działania produktu, programu lub usługi pochodzących od producenta innego niż IBM spoczywa na użytkowniku.

IBM może posiadać patenty lub złożone wnioski patentowe na towary i usługi, o których mowa w niniejszej publikacji. Przedstawienie niniejszej publikacji nie daje żadnych uprawnień licencyjnych do tychże patentów. Pisemne zapytania w sprawie licencji można przysyłać na adres:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
U.S.A.*

Zapytania dotyczące zestawów znaków dwubajtowych (DBCS) należy kierować do lokalnych działów własności intelektualnej IBM (IBM Intellectual Property Department) lub wysłać je na piśmie na adres:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokio 103-8510, Japonia*

INTERNATIONAL BUSINESS MACHINES CORPORATION DOSTARCZA TĘ PUBLIKACJĘ W STANIE, W JAKIM SIĘ ZNAJDUJE ("AS IS") BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (RĘKOJMIĘ RÓWNIEŻ WYŁĄCZA SIĘ), WYRAŹNYCH LUB DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU ORAZ GWARANCJI, ŻE PUBLIKACJA TA NIE NARUSZA PRAW OSÓB TRZECICH. Ustawodawstwa niektórych krajów nie dopuszczają zastrzeżeń dotyczących gwarancji wyraźnych lub domniemanych w odniesieniu do pewnych transakcji; w takiej sytuacji powyższe zdanie nie ma zastosowania.

Informacje zawarte w niniejszej publikacji mogą zawierać nieścisłości techniczne lub błędy drukarskie. Informacje te są okresowo aktualizowane, a zmiany te zostaną uwzględnione w kolejnych wydaniach tej publikacji. IBM zastrzega sobie prawo do wprowadzania ulepszeń i/lub zmian w produktach i/lub programach opisanych w tej publikacji w dowolnym czasie, bez wcześniejszego powiadomienia.

Wszelkie wzmianki w tej publikacji na temat stron internetowych innych podmiotów zostały wprowadzone wyłącznie dla wygody użytkownika i w żadnym wypadku nie stanowią zachęty do ich odwiedzania. Materiały dostępne na tych stronach nie są częścią materiałów opracowanych dla tego produktu IBM, a użytkownik korzysta z nich na własną odpowiedzialność.

IBM ma prawo do używania i rozpowszechniania informacji przysłanych przez użytkownika w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich autora.

Licencjobiorcy tego programu, którzy chcieliby uzyskać informacje na temat programu w celu: (i) wdrożenia wymiany informacji między niezależnie utworzonymi programami i innymi programami (łącznie z tym opisywanym) oraz (ii) wspólnego wykorzystywania wymienianych informacji, powinni skontaktować się z:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
U.S.A.*

Informacje takie mogą być udostępnione, o ile spełnione zostaną odpowiednie warunki, w tym, w niektórych przypadkach, zostanie uiszczona stosowna opłata.

Licencjonowany program opisany w niniejszej publikacji oraz wszystkie inne licencjonowane materiały dostępne dla tego programu są dostarczane przez IBM na warunkach określonych w Umowie IBM z Klientem, Międzynarodowej Umowie Licencyjnej IBM na Program lub w innych podobnych umowach zawartych między IBM i użytkownikami.

Dane dotyczące wydajności i cytowane przykłady zostały przedstawione jedynie w celu zobrazowania sytuacji. Faktyczne wyniki dotyczące wydajności mogą się różnić w zależności do konkretnych warunków konfiguracyjnych i operacyjnych.

Informacje dotyczące produktów innych podmiotów niż IBM zostały uzyskane od dostawców tych produktów, z ich publicznych ogłoszeń lub innych dostępnych publicznie źródeł. IBM nie testował tych produktów i nie może potwierdzić dokładności pomiarów wydajności, kompatybilności ani żadnych innych danych związanych z tymi produktami. Pytania dotyczące możliwości produktów innych podmiotów należy kierować do dostawców tych produktów.

Wszelkie stwierdzenia dotyczące przyszłych kierunków rozwoju i zamierzeń IBM mogą zostać zmienione lub wycofane bez powiadomienia.

Publikacja ta zawiera przykładowe dane i raporty używane w codziennej pracy. W celu kompleksowego ich zilustrowania, podane przykłady zawierają nazwiska osób prywatnych, nazwy przedsiębiorstw oraz nazwy produktów. Wszystkie te nazwy/nazwiska są fikcyjne i jakiegokolwiek podobieństwo do istniejących nazw/nazwisk jest całkowicie przypadkowe.

Znaki towarowe

IBM, logo IBM i ibm.com są znakami towarowymi lub zastrzeżonymi znakami towarowymi International Business Machines Corp. zarejestrowanymi w wielu systemach prawnych na całym świecie. Pozostałe nazwy produktów i usług mogą być znakami towarowymi IBM lub innych przedsiębiorstw. Aktualna lista znaków towarowych IBM dostępna jest w serwisie WWW IBM, w sekcji "Copyright and trademark information" (Informacje o prawach autorskich i znakach towarowych), pod adresem www.ibm.com/legal/copytrade.shtml.

Adobe, logo Adobe, PostScript oraz logo PostScript są znakami towarowymi lub zastrzeżonymi znakami towarowymi Adobe Systems Incorporated w Stanach Zjednoczonych i/lub w innych krajach.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium i Pentium są znakami towarowymi lub zastrzeżonymi znakami towarowymi Intel Corporation lub przedsiębiorstw podporządkowanych Intel Corporation w Stanach Zjednoczonych i w innych krajach.

Linux jest zastrzeżonym znakiem towarowym Linusa Torvaldsa w Stanach Zjednoczonych i/lub w innych krajach.

Microsoft, Windows, Windows NT oraz logo Windows są znakami towarowymi Microsoft Corporation w Stanach Zjednoczonych i/lub w innych krajach.

UNIX jest zastrzeżonym znakiem towarowym The Open Group w Stanach Zjednoczonych i/lub w innych krajach.

Java oraz wszystkie znaki towarowe i logo dotyczące języka Java są znakami towarowymi lub zastrzeżonymi znakami towarowymi Oracle i/lub przedsiębiorstw afiliowanych.

Warunki dotyczące dokumentacji produktu

Zezwolenie na korzystanie z tych publikacji jest przyznawane na poniższych warunkach.

Zakres stosowania

Niniejsze warunki stanowią uzupełnienie warunków używania serwisu WWW IBM.

Użytek osobisty

Użytkownik ma prawo kopiować te publikacje do własnego, niekomercyjnego użytku pod warunkiem zachowania wszelkich uwag dotyczących praw własności. Użytkownik nie ma prawa dystrybuować ani wyświetlać tych publikacji czy ich części, ani też wykonywać na ich podstawie prac pochodnych bez wyraźnej zgody IBM.

Użytek służbowy

Użytkownik ma prawo kopiować te publikacje, dystrybuować je i wyświetlać wyłącznie w ramach przedsiębiorstwa Użytkownika pod warunkiem zachowania wszelkich uwag dotyczących praw własności. Użytkownik nie ma prawa wykonywać na podstawie tych publikacji ani ich fragmentów prac pochodnych, kopiować ich, dystrybuować ani wyświetlać poza przedsiębiorstwem Użytkownika bez wyraźnej zgody IBM.

Prawa

Z wyjątkiem zezwoleń wyraźnie udzielonych w niniejszym dokumencie, nie udziela się jakichkolwiek innych zezwoleń, licencji ani praw, wyraźnych czy domniemanych, odnoszących się do tych publikacji czy jakichkolwiek informacji, danych, oprogramowania lub innej własności intelektualnej, o których mowa w niniejszym dokumencie.

IBM zastrzega sobie prawo do anulowania zezwolenia przyznanego w niniejszym dokumencie w każdej sytuacji, gdy, według uznania IBM, korzystanie z tych publikacji jest szkodliwe dla IBM lub jeśli IBM uzna, że warunki niniejszego dokumentu nie są przestrzegane.

Użytkownik ma prawo pobierać, eksportować lub reeksportować niniejsze informacje pod warunkiem zachowania bezwzględnej i pełnej zgodności z obowiązującym prawem i przepisami, w tym ze wszelkimi prawami i przepisami eksportowymi Stanów Zjednoczonych.

IBM NIE UDZIELA JAKICHKOLWIEK GWARANCJI, W TYM TAKŻE RĘKOJMI, DOTYCZĄCYCH TREŚCI TYCH PUBLIKACJI. PUBLIKACJE TE SĄ DOSTARCZANE W STANIE, W JAKIM SIĘ ZNAJDUJĄ ("AS-IS") BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (RĘKOJMIĘ RÓWNIEŻ WYŁĄCZA SIĘ), WYRAŹNYCH CZY DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU CZY NIENARUSZANIA PRAW OSÓB TRZECICH.

Glosariusz

A

wartość progowa dokładności . Kryterium zatrzymywania procesu przenikania odpowiadające najmniejszej zmianie energii węzła, która gwarantuje kontynuację przenikania.

lider autorytetu . Osoba w grupie, która ma najwyższy wynik autorytetu.

wynik autorytetu . Miara wskazująca częstotliwość, z jaką inni członkowie grupy łączą się z daną osobą. Wynik autorytetu kojarzy ważność osoby z liczbą relacji kończących się na tej osobie.

C

grupa kluczowa . Grupa węzłów pozostających po pominięciu słabszych relacji w sieci i zastosowaniu ograniczeń co do wielkości grupy.

wartość progowa pokrycia . Ilość najsilniejszych relacji wykorzystywanych do identyfikacji grupy.

D

stopień . Miara aktywności węzła w sieci, definiowana jako liczba relacji, w które zaangażowany jest ten węzeł.

gęstość . Miara spójności sieci, definiowana jako liczba zaobserwowanych relacji podzielona przez liczbę możliwych relacji.

relacja dychotomiczna . Powiązanie między dwoma węzłami; może przyjąć tylko jedną z dwóch wartości, wskazującą obecność lub nieobecność powiązania.

rozproszona energia . Wpływ pochodzący z określonych węzłów w sieci, który rozprzestrzenia się po całej sieci, zmniejszając się wraz z oddalaniem się od węzłów początkowych.

relacja kierunkowa . Powiązanie między dwoma węzłami, które rozpoczyna się w jednym (źródle) i kończy się w drugim (celu).

lider rozpowszechniania . Osoba w grupie, która ma najwyższy wynik rozpowszechniania.

wynik rozpowszechniania . Miara wskazująca częstotliwość łączenia się danej osoby z innymi osobami w grupie. Wynik rozpowszechniania kojarzy ważność osoby z liczbą relacji rozpoczynających się od tej osoby.

I

stopień wejściowy . Miara prestiżu węzła w sieci składającej się z ukierunkowanych relacji, definiowana jako łączna liczba relacji, w których dany węzeł jest celem.

N

relacja niekierunkowa . Związek między dwoma węzłami.

U

stopień wyjściowy . Miara centralności węzła w sieci składającej się z ukierunkowanych relacji, definiowana jako łączna liczba relacji, w których dany węzeł jest źródłem.

W

socjogram . Wizualne przedstawienie sieci społecznościowej, w którym linie łączące punkty obrazują osoby pozostające ze sobą w relacji.

współczynnik rozprzestrzeniania . Ilość energii, którą węzeł przekazuje do sąsiadujących węzłów w jednym kroku procesu przenikania.

D

relacja z wartościowaniem . Powiązanie między dwoma węzłami, z którym wiąże się waga wskazująca siłę relacji.

Indeks

A

- analiza grup 5
 - rozmiary grup 10, 12
 - wartość progowa pokrycia 10, 12
 - wymagania 11
- analiza przenikania 5, 15
 - opcje 16
 - wartość progowa dokładności 15, 17
 - współczynnik rozprzestrzeniania 15, 17
 - wymagania 16
 - ziarna 17
- analizowanie danych 6

C

- czyszczenie wyników 6

F

- filtrowanie danych wyjściowych 5

G

- GAG_GroupNumber 13
- gęstość 3, 12, 13
- gęstość jądrowa 12, 13

K

- kluczowe grupy 10
- kluczowe węzły 12, 13
- kluczowe wskaźniki wydajności 13, 18

L

- liderzy rozpowszechniania 10, 13
- liderzy w zakresie autorytetu 10, 13
- lista ziaren przenikania 17

P

- pola celu 7
- pola wagi 7, 11, 16
- pola źródeł 7
- poziomy pomiaru 5
- przeglądanie wyników 5

R

- rekordy dotyczące rozmów telefonicznych 7
- relacje
 - kierunek 1
 - z wartościowaniem 1
 - relacje dychotomiczne 1
 - relacje kierunkowe 1
 - relacje niekierunkowe 1
 - relacje z wartościowaniem 1

- rozmiary grup 12, 13
 - granice 10, 12
- rozproszona energia 17, 18

S

- skrypty
 - właściwości 7
- stopień 4
- stopień wejściowy 4, 12, 13, 17, 18
- stopień wyjściowy 4, 12, 13, 17, 18

Ś

- średni rozmiar grupy 12
- średni stopień wejściowy 12, 17
- średni stopień wyjściowy 12, 17
- średni wpływ 17
- średnia gęstość grupy 12
- średnia gęstość grupy kluczowej 12
- średnia ilość energii kluczowego członka 12

W

- wartość progowa dokładności 15, 17
- wartość progowa pokrycia 10, 12
- właściwości
 - skrypty 7
- współczynnik rozprzestrzeniania 15, 17
- wykonywanie strumieni
 - przy użyciu IBM SPSS Modeler Server Social Network Analysis 5
- wyniki autorytetu 10, 13
- wyniki rozpowszechniania 10, 13

Z

- zakładka adnotacje 5
- zakładka analiza 5
 - analiza grup 12
 - analiza przenikania 17
- zakładka dane 5
 - analiza grup 11
 - analiza przenikania 16
- zakładka filtr 5
- zakładka opcji kompilacji 5
 - analiza grup 12
 - analiza przenikania 17
- zakładka typy 5
 - ziarna
 - dla analizy przenikania 17



Drukowane w USA