

POLITECHNIKA WROCŁAWSKA
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI



Sieci złożone

Sprawozdanie z laboratorium

AUTOR

Alicja Dąbrowska

nr albumu: **260349**

kierunek: **Inżynieria systemów**

24 styczeń 2023

Streszczenie

Celem pracy jest badanie właściwości sieci złożonej, utworzonej na podstawie rzeczywistych danych. Analizie będzie poddawany międzynarodowy handel szczepionkami medycznymi. Praca obejmuje reprezentację danych w postaci grafu, wyznaczanie różnego rodzaju charakterystyk sieci, jak również detekcję społeczności w sieci przy wykorzystaniu metody maksymalizacji modularności (algorytm Clauset-Newman-Moore). W rezultacie dokonanej analizy udało się ustalić, że pandemia Covid-19 wpłynęła na rozwój sieci handlu międzynarodowego. Dodatkowo ustalone zostało, że w handel międzynarodowy angażuje się coraz więcej państw i jest on coraz bardziej zglobalizowany. Na podstawie algorytmu detekcji społeczności wykryte zostały trzy zgrupowania w sieci.

1 Wstęp – sformułowanie problemu

Handel międzynarodowy jest ważnym elementem gospodarki światowej, którego wydajność może przynieść wiele korzyści dla gospodarki kraju. Nie tylko wpływa on na zwiększenie ilości dostępnych dóbr w państwie, lecz także stymuluje do rozwoju zaplecza technologicznego państw. Należy jednak pamiętać, że handel międzynarodowy składa się z wielu złożonych relacji między różnymi państwami i każda lekka zmiana w jednej z nich, może odbić się na pozostałych krajach i ich relacjach. Z tego powodu handel międzynarodowy wymaga ciągłego monitorowania i analiz pod różnym kątem. W ramach pracy utworzona zostanie sieć globalnego handlu szczepionkami medycznymi między państwami w oparciu o bazę danych, udostępnionych na oficjalnej stronie WITS (World Integrated Trade Solution). Podjęta zostanie próba wywnioskowania z otrzymanej sieci informacji, jak zmieniła się jej struktura w latach 2017-2021 z uwzględnieniem wpływu pandemii Covid-19. Dodatkowo podjęta zostanie próba wywnioskowania, czy handel międzynarodowy staje się zglobalizowany, a także sprawdzone zostanie, czy sieć handlowa dzieli się na mniejsze społeczności.

2 Opis rozwiązania

2.1 Zbiór danych

Dane dotyczące handlu szczepionkami medycznymi między państwami zostały pobrane ze strony <http://wits.worldbank.org/>. Dane były udostępnione w formie pliku .csv. Baza została zapisana w postaci ramki danych biblioteki **Pandas**. Zawiera ona informacje o 21951 wymianach między państwami, o charakterze zarówno importu, jak i eksportu w latach 2017-2021.

2.2 Zastosowane podejście

Podejście zastosowane do rozwiązania problemu można przedstawić w czterech etapach:

I. Reprezentacja sieci handlu międzynarodowego

Pierwszym krokiem jest reprezentacja danych w postaci grafu skierowanego, w którym węzłami są państwa, natomiast połączeniami są wymiany handlowe między państwami. Połączenie z A do B oznacza państwo A eksportujące produkt do państwa B lub państwo B importujące z A. Waga krawędzi przedstawia wartość handlu (jednostka USD). W celu uproszczenia badań w dalszej części analizy użyty został graf nieskierowany.

II. Rozkłady węzłów i połączeń w sieci

Pierwszym badanym aspektem jest zmieniająca się struktura sieci, czyli to, jak zmienia się liczba połączeń między węzłami w sieci na przestrzeni lat, a także jaki jest rozkład nowo stworzonych połączeń w sieci. Kolejno zbadane zostaną rozkłady stopni wierzchołków sieci dla różnych lat. Znając właściwości sieci złożonej, należy się spodziewać, że rozkłady te będą potęgowe. Wymienione operacje są bardzo pomocne w określeniu nie tylko budowy sieci, ale także tego, czy handel międzynarodowy staje się bardziej zglobalizowany.

III. Miary centralności

Kolejnym krokiem jest badanie miar centralności sieci (stopień, pośrednictwo itp.), które oferują wgląd do tego, które z węzłów oraz połączeń są najważniejsze. Dzięki temu można zbadać, jakie państwa i połączenia handlowe są najistotniejsze w międzynarodowym handlu szczepionkami medycznymi oraz ocenić, jak ewoluuje sieć.

IV. Detekcja społeczności

Społeczności mówią wiele o strukturze sieci. Ich detekcja może wskazać, jakie są zależności między różnymi państwami. Do detekcji społeczności użyty został algorytm Clauset-Newman-Moore, oparty na modułowości. Algorytm zachłanny maksymalizacji modułowości rozpoczyna się od każdego węzła w jego własnej społeczności i wielokrotnie łączy parę społeczności, które prowadzą do największej modułowości, aż dalsze zwiększanie modułowości nie jest możliwe (maksimum).

3 Rezultaty obliczeń

3.1 Plan badań

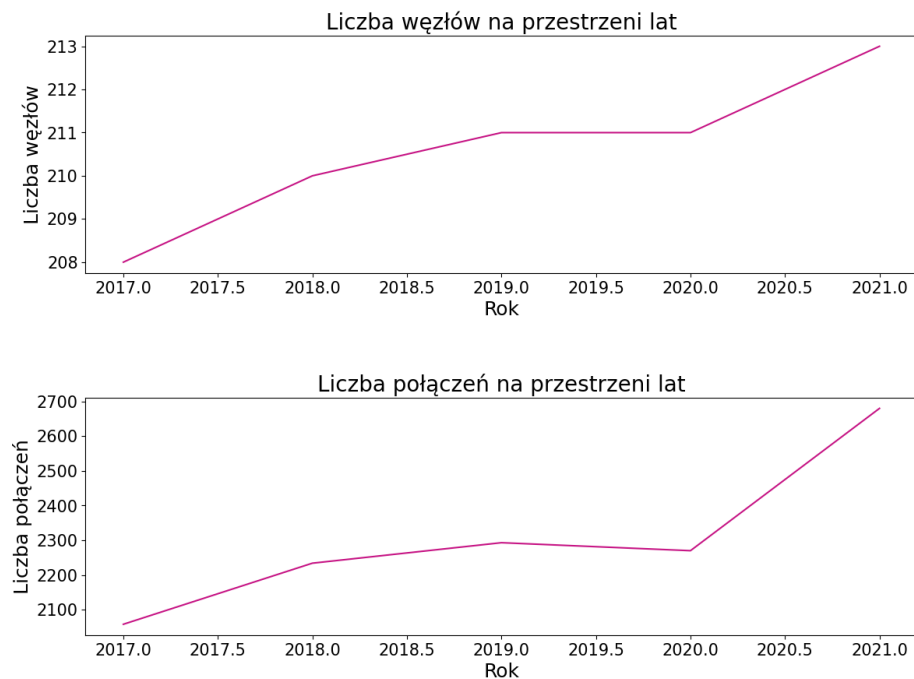
Do wczytania i transformacji danych skorzystano z metod biblioteki **Pandas**. W celu zamodelowania sieci oraz obliczenia charakterystyk sieci (stopnie, pośrednictwo) wykorzystany został pakiet **Networkx**. Detekcję społeczności przy wykorzystaniu maksymalizacji modułowości wykonano przy użyciu metody `greedy-modularity-communities(G)` z rozszerzenia `networkx.algorithms.community`.

3.2 Wyniki obliczeń

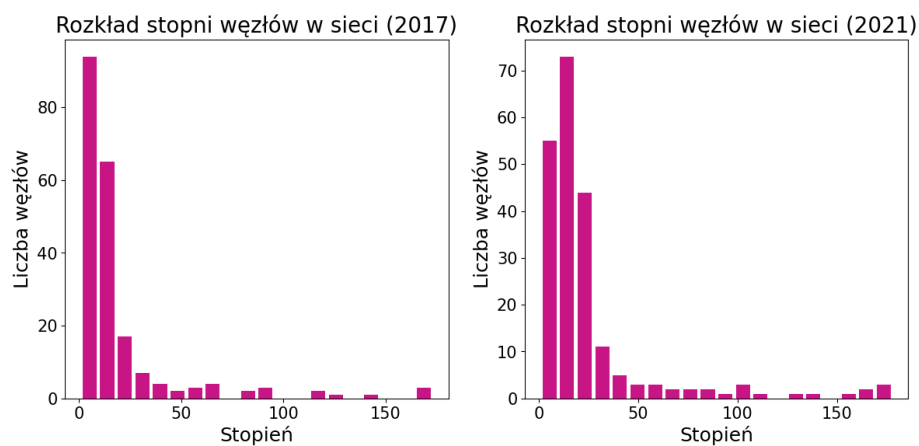
I. Globalne zmiany w sieci - analiza struktury

Analiza wyników rozpoczyna się od spojrzenia na zmianę liczby połączeń i węzłów na przestrzeni lat 2017-2021, co przedstawia rysunek 1. Można zauważyć, że zarówno liczba węzłów jak i połączeń stale rośnie, a dodatkowo w roku 2020 przyrost ten staje się jeszcze bardziej ostry. Może to być efektem wybuchu pandemii Covid-19, która przyczyniła się do wzrostu zapotrzebowania na szczepionki przez wiele państw. Wzrost węzłów i połączeń jest zrozumiałym efektem w sieci handlu międzynarodowego, ponieważ coraz więcej państw dołącza się do globalnych wymian handlowych, tworząc coraz więcej połączeń. Potwierdza to polityczną i ekonomiczną hipotezę, że handel międzynarodowy staje się coraz bardziej zglobalizowany. Bazując na wykresie i ogólnej wiedzy, można powiedzieć, że zjawisko globalizacji w tym przypadku zostało spotęgowane przez rozprzestrzenienie się pandemii.

Na rysunku 2 przedstawione są wykresy rozkładu stopni węzłów dla roku 2017 oraz 2021. Widać, że rozkłady mają charakter potęgowy, co jest podstawową cechą sieci złożonych. Oznacza to, że w sieci najwięcej jest węzłów o małej ilości połączeń, natomiast wraz ze wzrostem liczby połączeń,



Rysunek 1: Zmiana liczby węzłów i połączeń w sieci w latach 2017-2021.

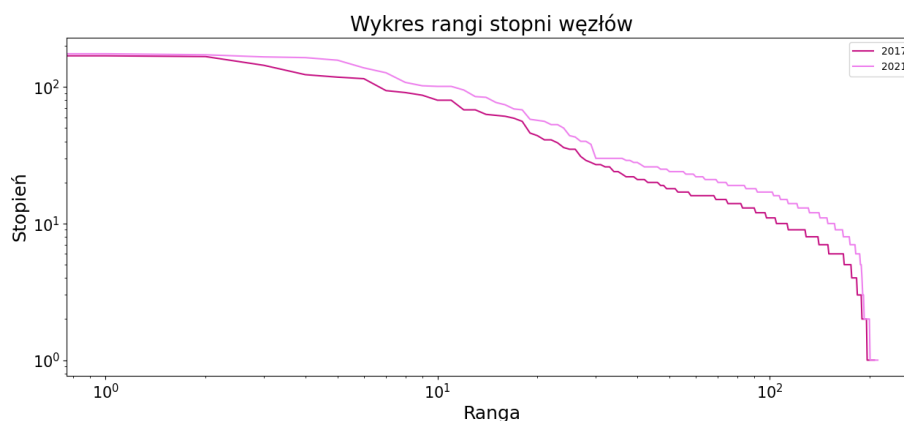


Rysunek 2: Rozkład stopni węzłów w roku 2017 oraz 2021.

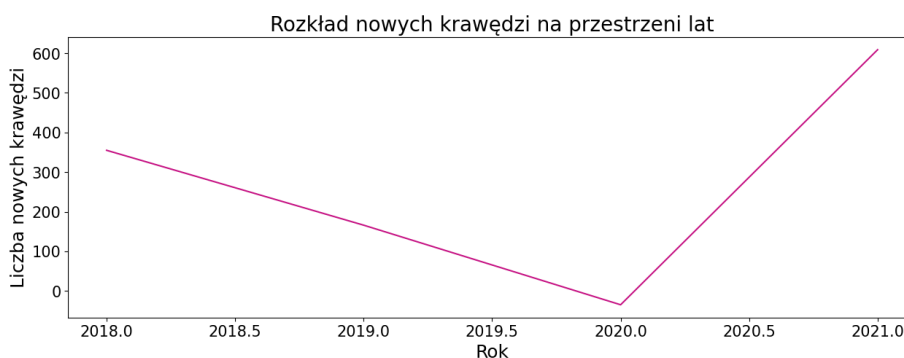
liczba takich węzłów szybko maleje. Można zauważyć, że w roku 2021 sieć ma lekko zaburzony rozkład potęgowej stopni wierzchołków w stosunku do sieci z roku 2021, ale także ma więcej węzłów o wyższych stopniach, natomiast zdecydowanie mniej węzłów o małych stopniach. Dodatkowo rysu-

nek 3 przedstawia wykres rangi stopni węzłów, który jest wykresem stopnia każdego węzła w sieci, posortowanym od najwyższego do najniższego, przedstawionym w skali logarytmicznej. Ranga to pozycja węzła na posortowanej liście węzłów według stopnia (ranga węzła o najwyższym stopniu to 1, drugiego najwyższego to 2 itd.). Z rysunku wynika, że stopnie wierzchołków w sieci z 2021r. rosną w stosunku do roku 2017. Powyższe obserwacje mogą wskazywać, że do handlu międzynarodowego przyłącza się coraz więcej państw oraz staje się on coraz bardziej rozwinięty, na co wskazuje pojawianie się większej ilości dobrze usieciowionych węzłów.

Rysunek 4 przedstawia dołączanie się nowych połączeń do sieci handlu międzynarodowego. Zauważyć można, że rok 2020 jest pewnego rodzaju odbiciem od wcześniejszego trendu rozkładu nowych połączeń. Taki wzrost liczby nowych krawędzi w roku 2020 może wynikać z wysokiego zapotrzebowania na szczepionki w czasie pandemii. Potwierdza to stwierdzenie, że wybuch wirusa Covid-19 przyczynia się do przyspieszenia rozwoju międzynarodowego handlu szczepionkami medycznymi.

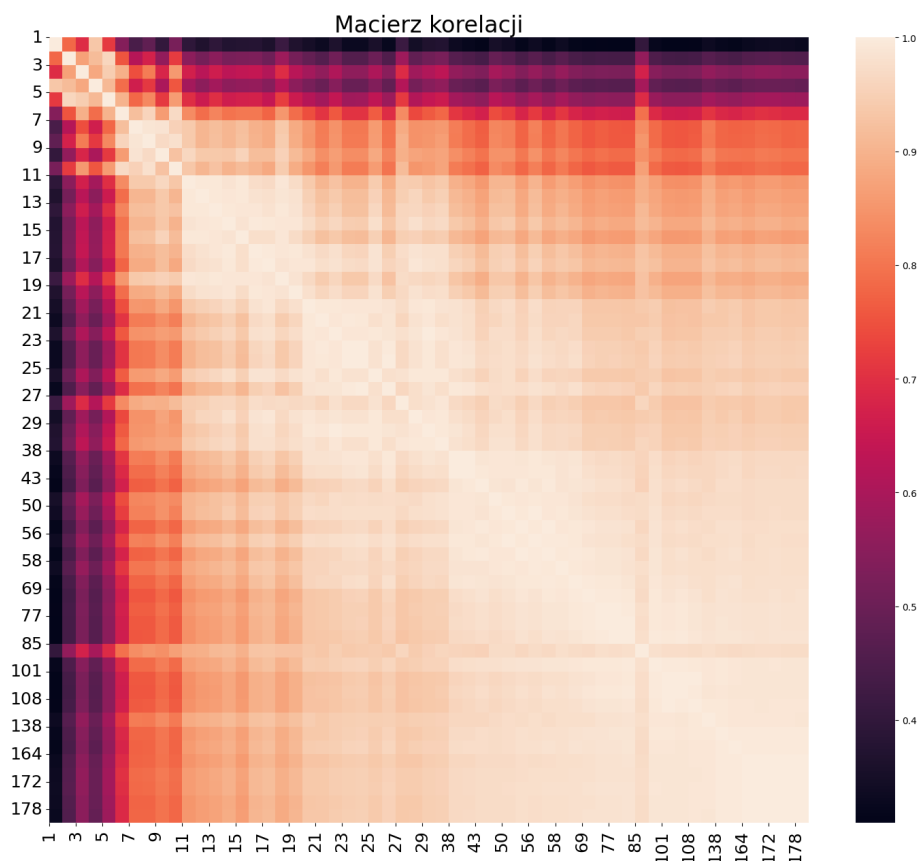


Rysunek 3: Wykres rangi stopni węzłów w roku 2017 oraz 2021.



Rysunek 4: Rozkład nowych krawędzi na przestrzeni lat 2017-2021.

Rysunek 5 przedstawia macierz korelacji między stopniami węzłów. Można zauważyć, że zdecydowanie wyższa korelacja występuje między węzłami o podobnym stopniu, co oznacza, że sieć ma charakter asortatywny. Można się zatem spodziewać, że bardziej prawdopodobne jest powstanie nowych połączeń między węzłami o tych samych lub podobnych stopniach.



Rysunek 5: Macierz korelacji między stopniami węzłów (2021)

II. Wyniki miar centralności

W ramach pracy wskazane zostały połączenia o najwyższym pośrednictwie, czyli liczbie lub odsetku najkrótszych dróg przechodzących przez krawędź. Tabela 1 przedstawia najważniejsze połączenia w roku 2017 na podstawie pośrednictwa. Widać, że handel jest mocno skierowany na Panamę, która pojawia się w większości najważniejszych krawędzi. Tabela 2 przedstawia sytuację dla roku 2021. Dominującą rolę w wymianie szczepionek medycznych między krajami pełnią Zjednoczone Emiraty Arabskie. Warto też wspomnieć, że wartość pośrednictwa krawędzi zdecydowanie wzrosła w stosunku do 2017, co może potwierdzać intuicję, że handel staje się coraz bardziej zglobalizowany.

Węzeł źródłowy	Węzeł docelowy	Pośrednictwo
Dominikana	Panama	96.7
Jamajka	Panama	47.3
Panama	Wyspy Brytyjskie	26.3
Haiti	Panama	22.5
Costa Rica	Panama	16.9
Honduras	Panama	10.2
Boliwia	Panama	9.1
Ekwador	Panama	8.3
Gujana	Panama	7.1
Bahamy	Panama	6.1

Tabela 1: Tabela przedstawia 10 najważniejszych krawędzi na podstawie pośrednictwa w roku 2017.

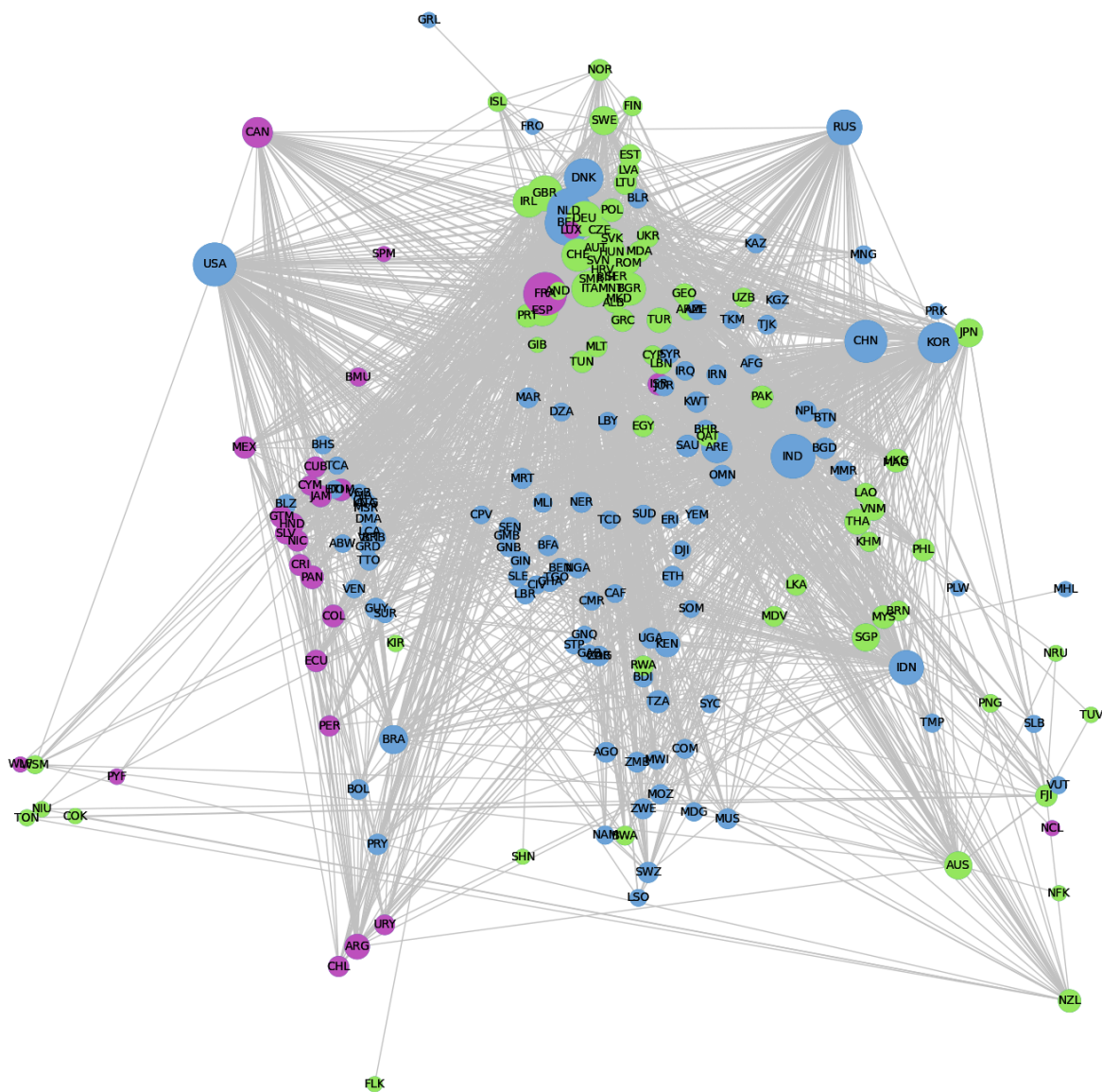
Węzeł źródłowy	Węzeł docelowy	Pośrednictwo
Zjednoczone Emiraty Arabskie	Sri Lanka	25346.4
Zjednoczone Emiraty Arabskie	Jemen	23329.7
Angola	Zjednoczone Emiraty Arabskie	22585.2
Zjednoczone Emiraty Arabskie	Bułgaria	14670.6
Zjednoczone Emiraty Arabskie	Brazylia	12464.3
Zjednoczone Emiraty Arabskie	Sudan	11822.8
Zjednoczone Emiraty Arabskie	Izrael	7254.9
Zjednoczone Emiraty Arabskie	Malezja	6361.3
Zjednoczone Emiraty Arabskie	Burkina Faso	6109.8
Zjednoczone Emiraty Arabskie	Terytorium Palestyny	6044.9

Tabela 2: Tabela przedstawia 10 najważniejszych krawędzi na podstawie pośrednictwa w roku 2021.

III. Detekcja społeczności

Zastosowany do detekcji społeczności algorytm, przyjmuję zasadę, że dobry podział grafu implikuje dużą liczbę krawędzi wewnątrz społeczności i małą liczbę krawędzi między społecznościami. W wyniku przeprowadzenia algorytmu doszło do podziału sieci na 3 społeczności, które prezentuje rysunek 6, z faktycznym położeniem geograficznym węzłów sieci. Pierwsza grupa (zielona) tworzy się, na poziomie Europy. Jednak również znajduje się w niej kilka państw terenów azjatyckich czy australijskich. Druga społeczność (fioletowa), obejmuje głównie teren Ameryki Południowej z dodatkami kilku państw spoza tamtego terenu. Największą społeczność (niebieską) tworzą tereny Afryki, Azji wraz z USA, które również zawierają w sobie węzły z wielu innych terenów. Przedstawiony podział sieci nie jest czysto geograficzny, jednak widać, że to, w jaki sposób węzły łączą się w społeczności, może być w jakimś stopniu uwarunkowane ich położeniem geograficznym.

Podsumowując, zbadana została zmieniająca się struktura sieci na przestrzeni lat, wskazująca, że wraz z upływem czasu wzrasta liczba połączeń i węzłów. Oznacza to, że sieć stale się rozwija, co dodatkowo mogło zostać spotęgowane przez wybuch pandemii Covid-19. Dokonana została także analiza asortatywności sieci, która wskazuje, że w sieci występują tendencja do tworzenia się nowych



Rysunek 6: Sieć handlu międzynarodowego szczepionkami medycznymi po detekcji społeczności

połączeń między węzłami o tych samych lub podobnych stopniach. Obliczone miary pośrednictwa krawędzi wskazują na ogólny wzrost pośrednictwa w stosunku do roku 2017. Wykonana detekcja społeczności wykazała 3 zgrupowania, które można wyodrębnić, bazując na modularności sieci.

4 Wnioski

Z otrzymanej sieci oraz jej analizy wywnioskować można informację, że międzynarodowy handel szczepionkami medycznymi jest siecią o stale zmieniającej się strukturze. Zdecydowanie zauważalne są zmiany w budowie sieci w roku 2020, co oznacza, że na międzynarodowy handel szczepionkami medycznymi duży wpływ mógł mieć wybuch pandemii Covid-19. Jednak w przeciwieństwie do wielu innych dziedzin wpłynął on pozytywnie na rozwój sieci, ze względu na duże zapotrzebowanie krajów na szczepionki medyczne. Z przeprowadzonej analizy widać również, że do handlu międzynarodowego dołącza się coraz więcej państw i staje się on coraz bardziej zglobalizowany. Po wykonaniu detekcji społeczności za pomocą algorytmu opartego na modularności, odkryte zostały 3 zgrupowania w sieci. Zatem przeprowadzone badania oraz analiza umożliwiły rozwiązanie problemu postawionego w punkcie 1.

5 Dodatek

Kod źródłowy został zamieszczony w repozytorium GitHub pod adresem: <https://github.com/alicjadabrowska2301/SZ-project>