

Pannon Egyetem
Műszaki Informatikai Kar
Kvantitatív Módszerek Intézeti Tanszék
Adattudomány MSc

DIPLOMADOLGOZAT

**Többrétegű kereskedelmi és kollaborációs hálózatok
jellemzőinek idősoros vizsgálata, anomáliák, sokkok
detektálása**

Pap Dávid

Témavezető: Dr. Kosztyán Zsolt Tibor

2025



PANNON EGYETEM

MŰSZAKI INFORMATIKAI KAR

Adattudomány MSc szak

Veszprém, 2025. április 2.

DIPLOMADOLGOZAT TÉMAKIÍRÁS

Pap Dávid

Adattudomány MSc szakos hallgató részére

Többrétegű kereskedelmi hálózatok jellemzőinek idősoros vizsgálata, anomáliák, sokkok detektálása

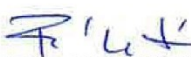
Témavezető: Dr. Kosztyán Zsolt Tibor, egyetemi tanár

A feladat leírása:

A kereskedelmi hálózatok vizsgálata már több mint tíz éves múltra tekint vissza. Ugyanakkor ezeket a vizsgálatokat elsősorban egy, vagy egymástól független rétegeken vizsgálták. A feladat olyan többrétegű hálózat kialakítása, illetve az élek kialakulásának gazdasági modellezése, amellyel magyarázható a rétegek közötti szereplők egymásra hatása is. A feladat során ki kell alakítani egy dinamikus többrétegű hálózati modellt. Ki kell számolni a hálózati és csomóponti mutatók idősoros adatait, majd ezt statisztikiai, gépi tanulási és ökonometriai módszerekkel elemezni is szükséges.

Feladatkiírás:

- Dolgozza fel a témakörhöz kapcsolódó szakirodalmat.
- Alakítson ki kereskedelmi forgalmakat tartalmazó adatokból többrétegű, dinamikus hálózatot.
- Számítsa ki és csoportosítsa a hálózati és csomóponti mutatók időbeli változásait.
- Értékelje és értelmezze az elemzés eredményeit.


Dr. Fogarassyné dr. Vathy Ágnes
egyetemi docens
szakfelelős


Dr. Kosztyán Zsolt Tibor
egyetemi tanár
témavezető

Hallgatói nyilatkozat

Alulírott Pap Dávid hallgató (Neptun kód: RJPEKF) kijelentem, és a dolgozat feltöltésével egyidejűleg nyilatkozom, hogy a Többrétegű kereskedelmi és kollaborációs hálózatok jellemzőinek idősoros vizsgálata, anomáliák, sokkok detektálása című diplomadolgozatot (a továbbiakban: dolgozat) a Pannon Egyetem Kvantitatív Módszerek Intézeti Tanszékén készítettem az okleveles adattudós oklevél megszerzése érdekében.

Kijelentem, hogy a dolgozatban csak a megadott és hivatkozott forrásokat használtam fel, és ezekre a vonatkozó idézési szabályok szerint hivatkoztam.

Nyilatkozom, hogy a dolgozat érdemi része saját szellemi alkotásom eredménye, és azt más intézményben, szakon, vagy felsőfokú képzés megszerzésére nem nyújtottam be. Tudomásul veszem, hogy a plágium vagy szerzői jogsértés esetén a dolgozatom elutasításra kerülhet, és ellenem fegyelmi eljárás indulhat. Tudomásul veszem továbbá, hogy szerzői jogsértés esetén az Egyetem jogosult a dolgozat elérhetőségét korlátozni, valamint eltávolítani a dokumentumot a dolgozatok tárolására szolgáló, a témát vezető szervezeti egység által meghatározott elektronikus zárt rendszerből.

Tudomásul veszem továbbá, hogy a Pannon Egyetem a dolgozat eredményeit saját céljaira eltérő írásbeli megállapodás hiányában a Pannon Egyetem Szellemi Tulajdon Kezelési Szabályzatában foglaltaknak megfelelően szabadon felhasználhatja.

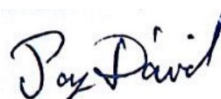
Nyilatkozom, hogy a dolgozat elkészítése során mesterséges intelligencia eszközöket használtam /nem használtam.

Nyilatkozom, hogy a dolgozat elkészítése során az alábbi táblázatban feltüntetett mesterséges intelligencia eszközöket kizárólag a kutatási, illetve fejlesztési feladat támogatására használtam fel, az érdemi munka, elemzés és következtetések teljes mértékben saját szellemi alkotásomat képezik.

Példa a táblázat kitöltésére:

Alkalmazott technológia	Alkalmazás módja	Előállított tartalom	MI használat aránya
Deepseek	Ötletelés, kutatási témák megállapítása	4. fejezet	30%
Deepseek	Szövegösszegzés	Tartalmi összefoglaló	20%
Gemini, ChatGPT	Kódgenerálás	Adattranszformációs műveletek	40%

Dátum: Veszprém, 2025. április. 24.



.....
Pap Dávid

Témavezetői nyilatkozat

Alulírott Dr. Kosztyán Zsolt Tibor témavezető kijelentem, hogy a Többrétegű kereskedelmi és kollaborációs hálózatok jellemzőinek idősoros vizsgálata, anomáliák, sokkok detektálása című dolgozatot Pap Dávid a Pannon Egyetem Kvantitatív Módszerek Intézeti Tanszékén készítette okleveles adattudós végzettség megszerzése érdekében. Kijelentem, hogy a dolgozat védelemre bocsátását engedélyezem.

Dátum: Veszprém, 2025. május. 3.



Dr. Kosztyán Zsolt Tibor

Köszönetnyilvánítás

Elsőként szeretném megköszönni családomnak és barátaimnak, akik támogattak engem azáltal, hogy mellettem álltak, biztattak és néha hasznos tanáccsal láttak el. Továbbá köszönettel tartozom témavezetőmnek, Dr. Kosztyán Zsolt Tibornak a segítségért és iránymutatásért.

Tartalmi összefoglaló

Diplomadolgozatomban elsősorban a globális kereskedelmi hálózatok elemzésére fókuszáltam, különös tekintettel a központi mutatókra, autoencoder modellekre, gráf neurális hálózatokra (GNN), embedding technikákra majd ezek segítségével az idősoros változások elemzésére és anomália-detektálásra.

Az elemzés alapját az OECD Inter-Country Input-Output (ICIO) adatbázis képezte, amely 1995 és 2020 között 76 ország kereskedelmi adatait tartalmazza, szektorokra bontva az ISIC Rev. 4 szabvány szerint. Ezeket az adatokat először adjacencia mátrixból él lista formátumba alakítottam, hogy aggregált és többrétegű hálózatok hozhassak létre. Az elkészült hálózatokból NetworkX gráfokat készítettem, majd a Gráf-neurális megoldások esetén PyGData objektummá alakítottam.

A készített gráfokon megvizsgáltam több központi mutatót, az így kapott eredményeket elemeztem, majd vizualizáltam és a kapott eredményeket dokumentáltam. Az adatok idősor elemzéséhez egy autoencoder alapú modellt és egy beágyazó modellt is készítettem. Az autoencoder alapú megoldás lehetővé tette az anomália-detektálást, vagyis a szokatlan vagy kiugró kereskedelmi mintázatok azonosítását a hálózatban amik gazdasági sokkokra utalhatnak.

Az embedding alapú megoldás segítségével egy idősoros vektorreprezentációt képeztem, amivel az évek során történt kereskedelmi hálózati átalakulásokat modelleztem. Majd ezeket a megközelítéseket a Gráf Neurális Hálózat segítségével is implementáltam, az így képzett GNN beágyazás a gráfok komplex kapcsolati struktúrájának tanulására használtam, amelyek képesek voltak a csomópontok és élek közötti nemlineáris összefüggések feltárására, így tovább javítva az anomáliák felismerését és az adathalmazban lévő hálózati dinamikák megértését.

Összességében a kutatás során központi mutatókat elemeztem, beágyazás segítségével időszori változásokat, autoencoderrel pedig anomáliákat detektáltam a kereskedelmi hálózatban. Majd ezekre készítettem egy Gráf Neurális alternatívát, amit kiértékeltem és összehasonlítottam a hagyományos módszerrel. A kapott eredmények lehetővé tették a hálózati dinamikák mélyebb megértését és a gazdasági rendellenességek pontosabb azonosítását.

Kulcsszavak: [Hálózatok, Központi mutatók, Autoencoder, Beágyazás, GNN]

Abstract

In my diploma thesis, I focused primarily on the analysis of global trade networks, with particular attention to centrality measures, autoencoder models, graph neural networks (GNN), embedding techniques, and with their help, the analysis of time series changes and anomaly detection.

The basis of the analysis was the OECD Inter-Country Input-Output (ICIO) database, which contains trade data for 76 countries between 1995 and 2020, divided into sectors according to the ISIC Rev. 4 standard. I first converted these data from an adjacency matrix to an edge list format in order to create aggregated and multilayer networks. I created NetworkX graphs from the completed networks, and then converted them into PyGData objects in the case of Graph-Neural solutions.

I examined several centrality measures on the created graphs, analyzed the results obtained in this way, then visualized and documented the obtained results.

For the time series analysis of the data, I created an autoencoder-based model and an embedding model. The autoencoder-based solution enabled anomaly detection, i.e. the identification of unusual or outlier trade patterns in the network that may indicate economic shocks.

Using the embedding-based solution, I created a time series vector representation, with which I modeled the trade network transformations that occurred over the years. I then implemented these approaches using the Graph Neural Network, using the GNN embedding formed this way to learn the complex connection structure of graphs, which were able to explore the nonlinear relationships between nodes and edges, thus further improving the detection of anomalies and the understanding of the network dynamics in the data set.

Overall, during the research, I analyzed centrality indicators, using embedding to detect time series changes, and using an autoencoder to detect anomalies in the trade network. I then developed a Graph Neural alternative to these, which I evaluated and compared with the traditional method. The results obtained allowed for a deeper understanding of network dynamics and a more accurate identification of economic anomalies.

Keywords: [Networks, Centrality measures, Autoencoder, Embedding, GNN]

Tartalomjegyzék

Jelölésjegyzék	10
1. Szakirodalmi áttekintés	11
1.1. Kereskedelmi hálózatelemzés.....	11
1.1.1. Adathalmazok.....	12
1.2. Központisági mutatók a hálózatelemzésben	12
1.3. Modern gépi tanulási módszerek a hálózatelemzésben	13
2. Adathalmaz kiválasztása	14
3. Központisági mutatók (Centrality Measures)	15
3.1. Fokszám központiság (Degree Centrality)	16
3.2. Közöttiség központiság (Betweenness centrality)	18
3.3. Közelségi központiság (Closeness Centrality)	20
3.3.1. Harmonikus központiság (Harmonic Centrality)	20
3.4. Sajátvektor központiság (Eigenvector Centrality).....	21
3.4.1. PageRank	21
3.5. Él köztesség (Edge Betweenness)	23
3.6. Denzitas (Density)	24
3.6.1. Rétegen belüli denzitas (Intra-layer density).....	26
3.6.2. Rétegek közti denzitas (Inter-layer density).....	28
3.7. Homofília (Homophily)	29
4. Közösségészlelés (Community Detection)	30
4.1. Kereskedelmi közösségek azonosítása	30
4.1.1. Louvain modularitás	30
4.1.2. Mag-periféria struktúrák.....	33
4.2. Közösségek idősor analízise	34
4.2.1. Dinamikus közösség detektálás (Dynamic Community Detection)	34
4.3. Anomália-észlelés.....	35
4.3.1. Autoencoder-alapú anomália-észlelés	36
4.4. kereskedelmi blokkváltások előrejelzése.....	39
4.4.1. Node2Vec és t-SNE.....	40
4.4.2. Eredmények kiértékelése	41

5. Gráf Neurális Hálók alkalmazhatósága	42
5.1. Gráf Neurális Hálózatok alkalmazása kereskedelmi hálózatokra	42
5.1.1. PygData	42
5.1.2. Konvolúciós rétegek	43
5.2. Kereskedelmi hálózatok beágyazása GNN segítségével	43
5.2.1. Algoritmus főbb lépései	43
5.2.2. A hálózat felépítése	43
5.2.3. Eredmények kiértékelése	44
5.3. Közösségészlelés	45
5.3.1. Motiváció	45
5.3.2. Eredmények kiértékelése	46
5.4. Anomália-észlelés	47
5.4.1. A hálózat felépítése	47
6. A programkódok elkészítése	50
6.1. GPU használat	50
6.2. Függvények tesztelése	51
6.2.1. Reprodukálhatósági teszt (Reproducibility testing)	51
6.2.2. Ablációs tesztelés (Ablation testing)	51
Irodalomjegyzék	52
Mellékletek	53
Ábrajegyzék	54
Táblázatjegyzék	55

Jelölésjegyzék

GNN:	Graph Neural Network
CEPII:	Centre d'Etudes Prospectives et d'Informations Internationales
WITS	World Integrated Trade Solution
OECD-	Organisation for Economic Co-operation and Development Inter-Country
ICIO	Input-Output
WTO	World Trade Organization
ROW	Rest of the World
ISIC	International Standard Industrial Classification
GNN	Gráf Neurális Hálózat

1. Szakirodalmi áttekintés

Ebben a részben bemutatom a kereskedelmi hálózatok elemzésének elméleti alapjait, a legfontosabb hálózatelemzési mutatókat, valamint a modern, gépi tanuláson alapuló módszerek alkalmazását a hálózatelemzésben. Az áttekintés bemutatja, hogy indult ki a gráfelemzésből a hálózatelemzés, hogy használják a központosági mutatókat és a különböző gépi tanulási módszereket hálózatanalízisre. Továbbá ezek az eszközök hogy teszik lehetővé a hálózatok szerkezetének, sérülékenységének és dinamikájának mélyebb megértését, ami kulcsfontosságú a globális gazdasági folyamatok elemzéséhez és a válságok hatásainak mérsékléséhez.

1.1. Kereskedelmi hálózatelemzés

A különböző hálózatelemzési technikák fontos részeit képezik a mindennapjainknak, legyen ez a közlekedés, internet, közösségi hálózatok vagy akár információ-áramlási kapcsolatok. Ezek tanulmányozása a gráfelméletből (Euler, 1736) és a 20. század eleji közösségi hálózatelemzésekből fejlődött ki, és az elmúlt két évtizedben a komplex hálózatok kutatása kiemelt jelentőségűvé vált. A komplex hálózatok olyan rendszerek, amelyek szerkezete szabálytalan, összetett és dinamikusan változik időben. Kutatók, mint Garlaschelli, Fagioli, De Benedict, Newman és Barabási, olyan hálózati struktúrákat vizsgáltak, ahol az egyes csomópontok (pl. egyének, közösségek) kapcsolatai meghatározzák a hálózat robusztusságát és reagálóképességét külső zavarokra, mint véletlenszerű hibák vagy célzott támadások. A tanulmányban ezekre alapozva gazdasági hálózatokra vizsgálja különböző módszerek alkalmazhatóságát és hatékonyságát. Egy lehetséges kutatási irány, a sokkelemzés is a hálózatelemzés sokat kutatott területe. Az ellátási láncokon keresztül ezek a sokkok hullámszerűen terjednek: ha egy régióban vagy iparágban visszaesik a termelés, az érinti a beszállítókat és a vevőket is, így a hatás az egész gazdaságra kiterjedhet, gyakran nagyobb mértékben, mint az eredeti sokk közvetlen hatása. Empirikus kutatások kimutatták, hogy például természeti katasztrófák vagy más váratlan események jelentős közvetett gazdasági hatásokat válthatnak ki az ellátási láncokon keresztül, különösen, ha a helyettesítő inputok nehezen elérhetők (Yasuyuki Todo, 2019). Az idézett irodalomban említésre kerül, hogy a sokkok mértéke és terjedése erősen függ a hálózati struktúra jellemzőitől, mint a sűrűség és a központi szereplők

jelenléte. A pénzügyi hálózatok esetében például a túlzottan sűrű kapcsolatok egy nagyobb sokk esetén felerősíthetik a hatásokat, míg kisebb sokkoknál a kevésbé sűrű hálózatok stabilabbak lehetnek. Ez egyáltalán azt is jelenti, hogy a hálózati szerkezet optimalizálása kulcsfontosságú a rendszer robusztusságának növelésében (Isabelle Rosenberg, 2018).

1.1.1. Adathalmazok

A globális kereskedelmi hálózatok vizsgálatára számos nagy, nyilvános adatbázis szolgál alapul, mint a UN Comtrade, BACI-CEPII, WITS, WTO International Trade Data és az OECD ICIO. Ezek közül az ICIO adatbázis különösen alkalmas a globális értékláncok és országok közötti, szektorokra bontott kereskedelmi kapcsolatok elemzésére, mivel részletes, időben követhető, 76 országot lefedő adatokat tartalmaz. A szakirodalom szerint az ilyen részletes, multilayer hálózatok elemzése elengedhetetlen a globalizáció, a kereskedelmi áramlások és a gazdasági sérülékenységek megértéséhez (Pieterse, 2015).

1.2. Központisági mutatók a hálózatelemzésben

A kereskedelmi hálózatok elemzésében a foksám központiság a közvetlen kapcsolatok számát méri, amely a kereskedelmi hálózatokban az exporttevékenység kiterjedtségét jelzi. A szakirodalom szerint a foksám központiság növekedése szoros összefüggésben áll a globalizációs trendekkel és a fejlődő országok felzárkózásával, hiszen a globalizáció előrehaladtával egyre több ország kapcsolódik be aktívan a nemzetközi kereskedelembe, és bővül az exportkapcsolatok száma (Kosztján et al., 2024). Ez a mutató jól tükrözi az országok kereskedelmi aktivitásának növekedését és a hálózat sűrűsödését.

A közöttiségi központiság azt mutatja, mely országok (vagy szektorok) töltenek be közvetítő szerepet a kereskedelmi hálózatban. A szakirodalom kiemeli, hogy gyakran a volt gyarmattartó nagyhatalmak, mint például az Egyesült Királyság vagy Franciaország, emelkednek ki ilyen közvetítőként, amit kulturális, történelmi és gazdasági örökségek magyaráznak (Szini, 2020; Vakhal, 2023).

A közelségi és harmonikus központiság mutatók a hálózat elérhetőségét és hatékonyságát mérik, így jól érzékeltetik az elszigetelt országok helyzetét. Különösen a keleti blokk és a fejlődő országok esetében ezek a mutatók segítenek feltárni, hogy milyen mértékben vannak beágyazódva a globális kereskedelmi hálózatba (Vakhal, 2023). A harmonikus

központiság előnye, hogy nem összefüggő gráfok esetén is alkalmazható, így szélesebb körben használható a kereskedelmi hálózatok elemzésében (Szini, 2020).

A sajátvektor központiság és a PageRank mutatók figyelembe veszik a szomszédos csomópontok jelentőségét, így komplexebb képet adnak a hálózatban elfoglalt szerepükről. A PageRank algoritmust eredetileg webgráfok elemzésére készítették, de mára elterjedt a kereskedelmi hálózatok vizsgálatában is. A szakirodalom rámutat, hogy a PageRank jól tükrözi a globális gazdasági erőviszonyok változásait, például Kína felemelkedését az USA rovására (Kosztján et al., 2024; Szini, 2020).

1.3. Modern gépi tanulási módszerek a hálózatelemzésben

Az utóbbi években egyre több kutatás alkalmaz mélytanulási technikákat, például autoencodereket és gráf neurális hálózatokat (GNN) a komplex hálózatok elemzésére és vizualizációjára. A beágyazási technikákat gyakran használják, amik képesek a magas dimenziójú hálózati adatokat „tömörített” reprezentációvá alakítani, amely megőrzi a hálózat szerkezeti sajátosságait. Ez lehetővé teszi a hálózatok vizuális feltérképezését és további gépi tanulási feladatok (klaszterezés, időszorelemzés) elvégzését (Hamilton et al., 2017). Az autoencoderek és gráf-autoencoderek alkalmazhatóságát magas hatékonyság jellemzi a gráf-elemzés esetén, ezért ez egy ajánlott technika a területen (Kipf & Welling, 2016).

Az ezekre a technikákra építő Gráf Neurális megoldások képesek a gráf szerkezetének és a csomópontok közötti bonyolult kapcsolatok megtanulására, ezzel pontosabb predikciókat és mintázatfelismerést tesznek lehetővé, mint a hagyományos módszerek. A szakirodalom szerint a GNN-ek különösen hatékonyak heterogén, multilayer hálózatokban, mint amilyenek a globális kereskedelmi hálózatok is (Wu et al., 2020).

A beágyazások és GNN-ek segítségével azonosíthatók a hálózatban szokatlan, kiugró mintázatok, amelyek gazdasági válságokra, szankciókra vagy strukturális változásokra utalhatnak. Ezt több tanulmány is alkalmazta pénzügyi hálózatokban (Akoglu et al., 2015), de kereskedelmi hálózatokra is egyre többen adaptálják (Panford-Quainoo et al., 2020).

2. Adathalmaz kiválasztása

Számos szervezet biztosít ingyenes hozzáférést a kereskedelmi hálózatokról tartott adatbázisaikhoz. Ezek közül néhány:

- **UN Comtrade Database** - Egy megbízhatónak és pontosnak tartott adatbázis, amiben részletes kategorizálással szerepelnek a termékek és szolgáltatások egyaránt, kereskedelmi érték szerint.
- **BACI-CEPII** - Az adatbázisok közül a UN és a CEPI adathalmazában vannak a termékcsoportok a legrészletesebben elkülönítve (az ISIC Rev. 4-nek megfelelően). A BACI-nak az imént ismertetett UN Comtrade adatbázis az alapja, amin különböző változtatásokat és javításokat alkalmaztak a fejlesztők.
- **WITS** - A worldbank oldalán elérhető kereskedelmi statisztikák. A fő hátránya az adathalmaznak, hogy a kereskedelmi útvonalakat összegzi nagyobb termékcsoportokba. Azonban ez az adathalmaz rendelkezik a legtöbb indikátorral, illetve a kereskedelmi összesítés havi nézetben is megtekinthető.
- **WTO International Trade data** - A többi adathalmazzal szemben a World Trade Organization a tarifákat is feltünteti az adathalmazban, továbbá külön szerepelnek a szolgáltatásokból származó kereskedelmi tevékenységek. Azonban sajnos ez nem tartalmaz részletes, országpáronkénti bontást.
- **ICIO** - Az ICIO-nak (Inter-Country Inter-Output) két változata létezik, egy általános és egy kiterjesztett. Az általános változat 76 ország adatait tartalmazza és egy ROW (rest of the world) oszlopot, míg a kiterjesztett Kína és Mexikó vonatkozásában részletesebb képet mutat.

A globális gazdaságot évtizedek óta a globalizáció, a globális értékláncok felemelkedése és a megnövekedett kereskedelmi áramlások alakították. Habár a globalizáció a geopolitikai feszültségek miatt lelassult, illetve a világjárvány megmutatta, hogy az ellátási láncok milyen sérülékenyek lehetnek, a globális értékláncok megértése továbbra is kulcsfontosságú.

Az OECD Inter-Country Input-Output (ICIO) adatbázisa infrastruktúrát biztosít az említett kérdések megválaszolásához az országokon belüli termelés, fogyasztás, befektetések és az országok közötti nemzetközi kereskedelem feltérképezésével. Ezek az adatok elengedhetetlenek az országok által a globális értékláncokban való részvétel révén generált hozzáadott érték megértéséhez. Továbbá az ICIO adathalmaz használatával

lehetőség nyílik a szektorok közötti áramlás vizsgálatára, ezeken mutatók számítására, illetve az adathalmazból készült gráfot több szempontból lehet többretegű hálózattá alakítani. A dolgozatomban ezért döntöttem az ICIO adathalmaz használata mellett.

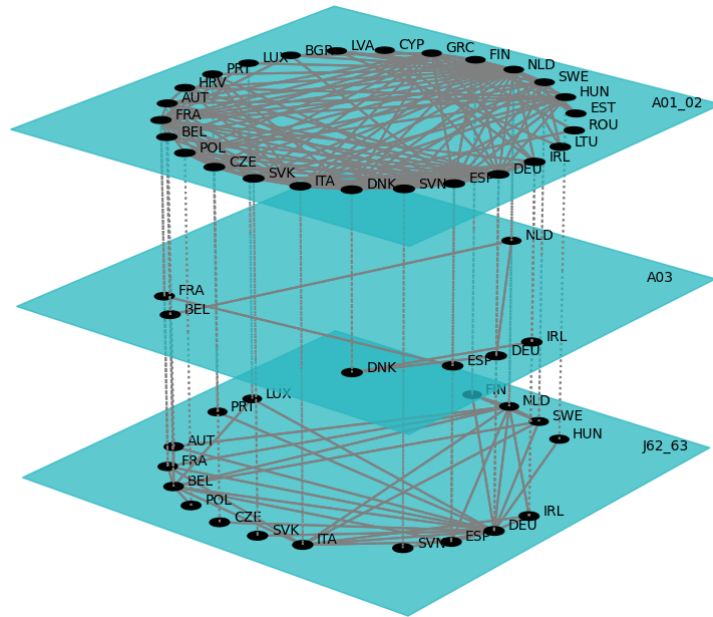
Az adathalmaz 1995-től 2020-ig periodikusan (évenkénti rögzítésében) tartalmaz adatokat, amik az imént említett 76 országról ISIC (International Standard Industrial Classification) 4-es sztenderd alapján tartalmaz kategorizált leírást.

Az adathalmazt későbbi könnyebb felhasználás céljából átalakítottam adjacencia mátrix formátumról él lista formátumra, az ország és szektorai elkülönítésével.

3. Központisági mutatók (Centrality Measures)

A központisági mutatók megmutatják a csomópontok szerepét vagy fontosságát egy hálózatban. Ezek a mutatók továbbá országok vagy szektorok szerepeit és befolyását is képesek számszerűsíteni.

Az elemzésekhez a többretegű hálózatok kialakítása többféle módon történt. Az első esetben a rétegek közötti kapcsolatok eredeti (ICIO) formájukban alakítottam gráffá, ahol a csúcsok országok szektorai az élek pedig az országok szektorai között adott évben létrejött kereskedelem mennyisége. Majd az így létrejött hálózatot a rétegek (szektorok) szerint aggregáltam így csak a különböző országok közötti kereskedelmet láthatóvá téve. Majd országok szerint is elvégeztem az aggregációt. Egyes esetekben pedig a hálózatot inter, majd intra-layer szempontból is elemeztem. A hálózatot multilayer hálózatként értelmeztem ahol a rétegek országok vagy szektorok voltak aggregálás nélkül. Ezeknél a hálózatoknál a rétegbeli (intra) vagy rétegeközi (inter) kapcsolatok elhagyásával végeztem elemzéseket, amik így diszjunkt gráfokat képeztek.



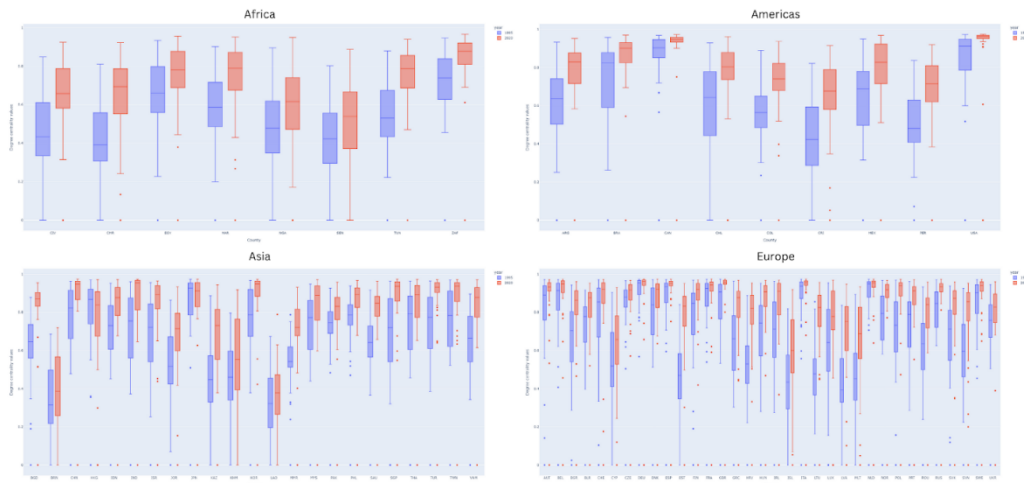
ábra 1: Intralayer vizualizáció, az EU országairól három szektor kereskedelmével

A példa ábrán a láthatóság érdekében csak az Európai unió országai látszanak, és azok 3 szektorban létrejött kereskedelmi útvonalai, ami meghaladják az 1 millió USD kereskedelmet.

A három ábrázolt szektor az Erdőgazdálkodás és mezőgazdasági (A01_02), a Halászati (A03) és a Programozással, adatműveletekkel (J62_63) kapcsolatos tevékenységek.

3.1. Fokszám központiség (Degree Centrality)

A központisági mutatók közül a legegyszerűbb a fokszám központiség, ami azt mutatja meg, hogy egy csomópontnak hány közvetlen kapcsolata van. Az ICIO adatahalmazból épített gráf irányított gráf, ezért ki-fokszám központiség és be-fok központiség is megállapítható, az elemzésben ki-fokszám központiséget használtam, ami az országok export tevékenységének kiterjedését hivatott reprezentálni. Azonban ez a mutató sajnos mivel nem tesz különbséget a csúcsok közt a kereskedelem mennyiségét illetően, csak a csúcspontok összekötöttségét tudja mérni. A mutató értéke normált, értéke így elosztva a maximum lehetséges élek számával $[0;(n-1)]$ közötti értéket vesz fel.



ábra 2: Kontinensek szerinti fokszámközpontiság érték 1995-től (kék) 2020-ig (piros)

A grafikonon jól látszik, hogy az évek során a kereskedelem növekvő szerepe révén a szektorok között egyre több kereskedelmi útvonal létesült ezáltal egyre több direkt kapcsolat jött létre. Míg 1995-ben a fokszám központiságok átlagos értéke 0.6636 volt, addig ez 2020-ra 0.7924-ra nőtt. A legkevesebb közvetlen összeköttetéssel rendelkező szektorok 1995-től 2020-ig Brunei (0.3394 -> 0.3812), Szenegál (0.3337 -> 0.5143) és Laosz (0.4323 -> 0.3718) voltak.

Észtország (0.4469 -> 0.7710) habár kis fokszám központiságról indult, mára felzárkózott az átlaghoz. Brunei, Laosz és Szenegál esetén ezt az egyenlőtlenséget a fejlődő és fejlett országok közötti kereskedelmi struktúrák különbségei, illetve az egyes országok gazdasági fejlettségi szintje is magyarázza (Pieterse, 2015).

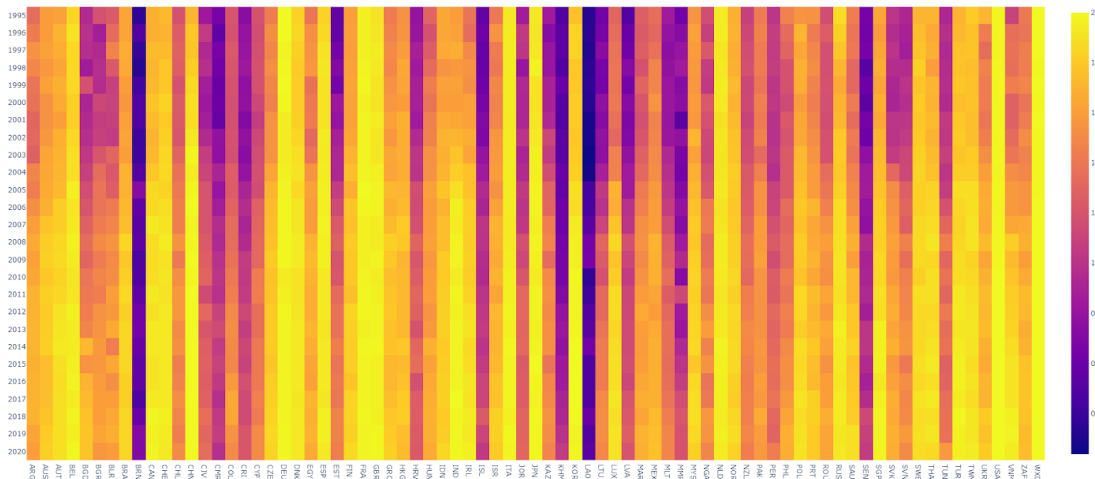
Az aggregált hálózati elemzés esetén csak az 1 millió dollárnál nagyobb kereskedelmű útvonalakat vettem figyelembe, mert minden útvonal figyelembe vétele esetén egy teljes gráf keletkezett.

A szektorok által aggregált hálózatra évenként általában ez az érték csak a B09 (bányászati segédszolgáltatás) a C30 (Egyéb szállítóeszközök gyártása) és a H53 (Posta és futárszolgálat) esetén volt kevesebb az érték, a többi szektor direkt módon csatlakozik minden másik szektorhoz.

Az országok által aggregált aggregált hálózat esetén az évek során látható, hogy az országok közti kereskedelem bővülő tendenciát mutat, azaz egyre bővült az országok kereskedelme. A '90-es években az észak amerikai, a nyugat európai országoknak, Japánnak és Dél-Koreának volt a legkiterjedtebb kereskedelmi hálózata. A legnagyobb

javulás Vietnám esetén történt, itt az 1 millió USD-t meghaladó kereskedelmi kapcsolatok száma majdnem megduplázódott.

Az ábrán látható egy éles törésvonal 2004 és 2005 között, ami fakadhat abból, hogy 2005-ben Kína felértékelte a valutáját a jüant, ami elsőre nem tűnhet jelentősnek. Azonban egyrészt növelte a kínai importigényt, ami serkentette a világkereskedelmet. Másrészt csökkentette a kereskedelmi feszültségeket, és hozzájárult a világgazdaság kiegyensúlyozottabbá tételéhez.



ábra 3: Országokénti Fokszám központosság 1995-től 2020-ig

Az ábrán továbbá kimutatható a 2008-as gazdasági világválság, amit a világosabb 2008-as sort egy annál sötétebb árnyalatú 2009-es követi. Ebből következik, hogy a válság hatására az 1 millió Dollárnál nagyobb forgalmú kereskedelmi hálózatok száma csökkent a világkereskedelemben, azon belül is leginkább a fejlődő országokat érintve.

3.2. Közöttiség központosság (Betweenness centrality)

A közöttiség központosság egy olyan mutató, ami azt méri, hogy egy csúcs milyen mértékben fekszik más csúcsok közötti utakon. Az ICIO adathalmazon az ország azon szektorai, amik magas közöttiség központissággal rendelkeznek gyakran közvetítőként kritikus fontosságú szerepet töltenek be.

Az alábbi grafikon azt mutatja, hogy kik voltak a jelentős közvetítő szerepet betöltő országok. Ezek az országok nagyrészt gyarmatosító nagyhatalmak voltak, és a volt gyarmataikkal kialakult szoros kapcsolatuk miatt ők szolgáltak közvetítőként a kereskedelmi kapcsolatokban. Ez nagyrészt kulturális, nyelvi és a gazdasági rendszerek

örökségéből adódik, ami azt jelenti, hogy sok esetben a gyarmatok általában főleg nyersanyagokat biztosítottak, az anyaországok pedig ipari késztermékeket.



ábra 4: Közöttiségi központisági értékek 1995-től 2020-ig gyarmatosító országok esetében.

Azonban ez nem minden esetben volt így, erre kitűnő példa Portugália, ahol a dekolonizáció csak későn történt meg. Esetükben a történelmi sérelmek, egyéb politikai feszültségek és a gazdasági érdekkülönbségek akadályozhatták a kereskedelmi együttműködést.

A gazdasági leigázás (economic subjugation) egy olyan helyzet, ahol egy erősebb fél egyenlőtlen feltételekkel (a másik fél számára kedvezőtlen kereskedelmi megállapodásokkal, adósságcsapdákkal, etc.) kényszerít egy másik felet a függésre, korlátozva annak gazdasági fejlődését és függetlenségét ez az Egyesült Államok esetén volt például jellemző (lásd: United Fruit Company).

Továbbá az ábrán látszik, hogy a mutatók értéke csökken, ami annak köszönhető, hogy a globalizáció miatt egyre több ország létesít egymással direkt kereskedelmi kapcsolatokat így a közvetítő országok egyre csekélyebb szerepet töltenek be.

A gazdasági szankciók és embargók is megfigyelhetők ezzel a mutatóval, ugyanis a szankciókkal sújtott országok szektorai és a büntető ország közötti kereskedelem csökken, vagy teljesen megszűnik, ezáltal csökken a csúcs közöttiségi központisági értéke. Azonban mivel az adathalmaz az ebben az időszakban jelentős szankciókkal sújtott országok adatait csak egyéb kategóriába sorolva tartalmazza, ezeket nem tudtam elemezni.

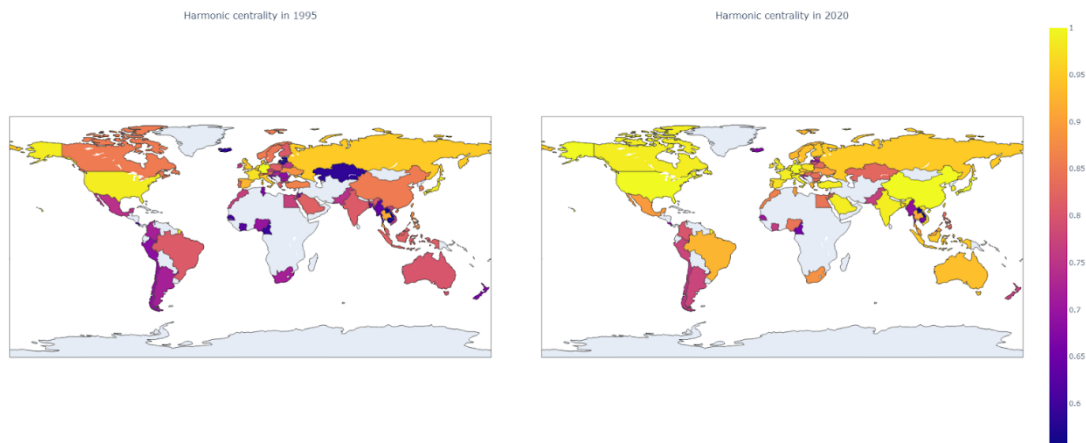
3.3. Közelségi központiség (Closeness Centrality)

A közelségi központiség azt méri, hogy hány lépésre van szükség ahhoz, hogy egy adott csúcsból minden más csúcsot elérjünk. Esetünkben ez kitűnő arra, hogy a közvetett elérést vizsgáljuk, adott ország szektorából a többi ország más szektorai hány lépéssel érhetőek el. Az alacsony lépésszám a kereskedelmi forgalom szempontjából hatékony útvonalakat és jó összekapcsoltságot jelent. A foksám központiségi mutató csak a közvetlen kapcsolatok számát képes mérni, még a közelségi központiségi mutató már a pont nagyobb környezetére is kiterjed, ezzel több csúcs elérését vizsgálva.

3.3.1. Harmonikus központiség (Harmonic Centrality)

A harmonikus központiség a közelségi központiség egy módosított változata, amely nem összefüggő gráfok esetében is jól alkalmazható. Esetünkben ez azért előnyös, mert az elhanyagolhatóan kis kereskedelmű utak elhagyásával izolált pontok, vagy egymással nem összefüggő komponensek keletkezhetnek.

A következő ábrán az országok szerinti aggregált Harmonikus központiségi mutató értékek láthatók.



ábra 5: Harmonikus központiségi értékek (térképen) 1995-ban és 2020-ban

A foksám központiséghez hasonlóan itt is megfigyelhető a keleti blokk országainak az elszigeteltsége, azonban a harmonikus centralitás esetén ez jobban szembetűnő, ugyanis itt a kereskedelmi partnerek kereskedelmi partnerei is befolyásolják a mutató értékét. Így mivel ezek az országok főleg egymással kereskedtek a harmonikus központiségi értékeik is alacsonyabbak. Ez a balti államok esetén látszik a leginkább, ahol az átlag 0.56, ami Laosz és Kambodzsa előtt a legalacsonyabb.

A harmonikus központiság és a következő fejezetben ismertetett PageRank közti fő különbség, hogy PageRank esetén a csúcs jelentősége a szomszédos csúcsok PR értékétől függ, míg harmonikus központiság esetén a szomszédos csúcsok (és azok szomszédai) számától függ és ezek távolságától. Esetünkben ez a távolság kereskedelmi forgalmat jelent, Dollárban kifejezve.

3.4. Sajátvektor központiság (Eigenvector Centrality)

A sajátvektor központiság hasonló a fokszám központisághoz, azonban ez a mutató nem csupán a csúcspontokhoz kapcsolódó élek számát veszi alapul, hanem azok “minőségeit” is. Ez azt jelenti, hogy amennyiben egy csúcs magas sajátvektor központiság értékű szomszédokkal rendelkezik, az ő sajátvektor központisági értéke is magas lesz. Ebből adódóan, egy csúcs rendelkezhet sok szomszéddal, azonban ha ezek kevésbé fontos csomópontok az eredeti csúcsunk értéke is alacsony lesz. A sajátvektor központiságot általában nem irányított gráfok esetén célszerű használni, mert kevésbé erősen összefüggő gráfok esetén a kevés összeköttetéssel rendelkező csúcspontok körül az eredmény torzul. Esetünkben a sajátvektor központiság egy változata, a PageRank sokkal jobb választás, amit a következő pontban ismertetek.

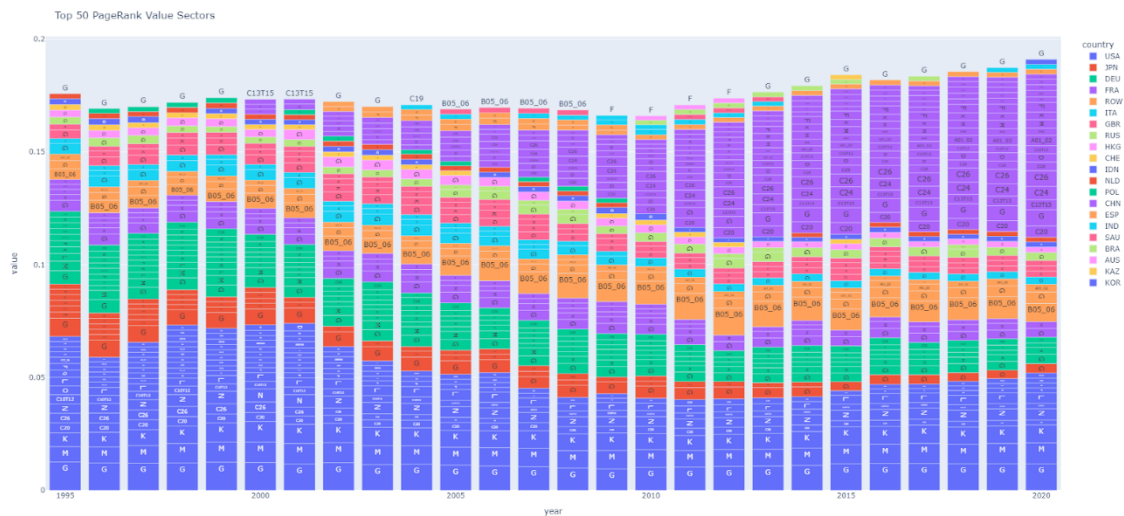
3.4.1. PageRank

A PageRank egy Google által fejlesztett algoritmus, ami a weboldalak fontosságuk alapján rangsorolja. Az eredeti szabadalom a Google egyik alapítójától, Larry Page-től származik és a keresőmotor sikerét és mai monopol helyzetét részben ennek a szabadalomnak köszönheti.

Működését tekintve az algoritmus bemenetét egy irányított gráf képezi, ahol a csúcsok a weboldalak az őket összekötő élek pedig a hiperhivatkozások jelölik. Az algoritmus először inicializálja a hálózat PR értékeit, majd iteratívan frissíti őket a szomszéd csúcsok értékeit figyelembe véve, míg el nem éri a kívánt iteráció számot vagy nem konvergál megfelelő mértékben a stabil értékhez. Egy úgynevezett csillapítási (damping) faktort is használ, hogy a modulok, amikből nem mutat ki él, ne torzítsák az eredményt. Az algoritmus hasonló módon működik a manapság igen elterjedt figyelmi (attention) mechanizmushoz (pl: LLM transzformerek).

A PageRank először csak webgráfokra, majd közösségi hálókra és nemrég kereskedelmi hálók elemzésére is gyakran kerül alkalmazásra. Míg a webgráfok és közösségi hálók élei általában súlyozatlanok ([uni-pannon.hu] - hipervivatkozás -> [facebook.com], [Alice] - ismeri -> [Bob]), addig a kereskedelmi hálók élei súlyozottak és ez a kereskedelem értékét hivatott kifejezni.

A következő ábra a PageRank által meghatározott 50 legfontosabb kereskedelmi szektort hivatott szemléltetni időrendben 1995-től 2020-ig. A szeletek mérete a szektor jelentőségét reprezentálja, pontosabban a meghatározott PageRank értéket.



ábra 6: Top 50 Pagerank értékű szektor 1995-től és 2020-ig országok által csoportosítva

Az évek során a legjelentősebb országok az Egyesült Államok, Kína, Németország, Egyesült Királyság, Franciaország, Olaszország és Oroszország. A grafikonon megfigyelhető, ahogy az Amerikai szektorok elvesztik dominanciájukat és Kína prominens szereplővé válik. Míg 1995-ben még csak két szektor, [G]Nagy- és kiskereskedelem, járműipar és a [C13T15]Textilipar számított világszinten jelentősnek, ez 2020-ra 21 szektorra nőtt és az imént említett gépjármű és textil szektorok is jelentős arányban nőttek. Érdekesség továbbá, hogy a [G]Nagy- és kiskereskedelem általában az országok legfontosabb csomópontja és általában itt is folyik a legjelentősebb kereskedelem, ez a trend az idősorban végig megfigyelhető.

Az 1995-től 2020-ig viszonylag változatlan helyzetű szektorok:

- [G]Nagy- és kiskereskedelem, járműipar
- [M]Szakmai, tudományos és műszaki tevékenység
- [L]Ingatlankereskedelem

-
- [C10T12]Élelmiszeripar, italok és dohánytermékek
 - [F]Építőipar
 - [K]Pénzügyi tevékenység, biztosítás
 - [N]Adminisztratív és támogató szolgáltatás

Az építőipart, ingatlankereskedelmet és az élelmiszeripart általában az országok vámokkal, kereskedelmi korlátozásokkal és specifikus korlátozásokkal védik a külföldi versenytársakkal szemben. Az építőipari beruházásokra pedig jellemző, hogy az államok a hazai vállalkozásokat részesítik előnyben külföldi versenytársaikkal szemben. Az élelmiszeriparban szigorú szabályozásoknak kell eleget tenni a kereskedelmi partnereknek. Az amerikai FDA (Food and Drug Administration) vagy az európai EFSA (European Food Safety Authority) korlátozásainak jelentős külföldi potenciális áruforgalom nem felel meg. Az ingatlankereskedelem az élelmiszeriparhoz hasonlóan is általában állami védelem alatt áll, külföldi entitások befektetési céllal kevés vagy egyáltalán nem vehetnek ingatlant. Habár ilyen erőteljes védelem alatt nem áll, de a pénzügyi szolgáltatások, biztosítások is általában országhatáron belül történnek.

Azonban a következő szektorokban figyelhető meg kirívó átrendezkedés:

- [C20]Vegyipar
- [C24]Fémipar
- [C26]Számítógép és elektronikai ipar

Az olcsó munkaerőnek és a bőséges erőforrásoknak köszönhetően Kína az említett iparágakban kulcsfontosságú szereplő lett, ezzel visszaszorítva versenytársait.

3.5. Él közteség (Edge Betweenness)

Hasonlóan a közöttségi központisághoz, az él közteségi érték egy adott élről kiszámítja, hogy az milyen gyakran fekszik a legrövidebb úton két csúcspont között. Az ICIO adathalmaz esetén a kereskedelmi utak fontosságát mutatja meg az él közteség és hogy mely kereskedelmi kapcsolatok kritikusak a globális kereskedelemben. A kritikus esetünkben azt jelenti, hogy kiesésük esetén, mivel ezek jellemzően alacsony fokszámú csúcsokat kötnek össze, ezek a csúcspontok egy útvonal kiesésével sokkal kitettebbek és sérülékenyebbek.

Az él közteségi értékek kiszámítása után a 10 legmagasabb értékkel rendelkező él mind Brunei gyógyszeripari, botanikai termékeivel és egy másik ország szektorával (legtöbb

esetben bányai par) áll összeköttetésben. Ez valószínűleg Brunei kereskedelmi elszigetelődéséből adódik (ahogy a foksám központiság résznél megállapítottam). Ezért Brunei-t elszigeteltsége miatt jobban érintené egy adott kereskedelmi kapcsolat elvesztése, különösen a gyógyszeriparát.

3.6. Denzitás (Density)

A denzitás az ICIO adathalmaz esetén a globális értékláncok (GVC - Global Value Chain) évenkénti alakulásába ad betekintést. Esetünkben az alacsony denzitás kevésbé összekapcsolt hálózatot jelent, míg az ellentétes sűrű, magas denzitás értékkel rendelkező hálózatokat jelentenek, amiknél magasabb a kohézió és a csomópontok összefüggőbbek. Emellett denzitás érték változása a globális szintű gazdaságpolitikai hatásokat is képes kimutatni. Például a denzitás értékének csökkenése jelezheti a globális kereskedelmi aktivitás visszaesését, amelyet gazdasági válságok, protekcionista intézkedések, vagy geopolitikai konfliktusok idézhetnek elő (országok közti kereskedelmi kapcsolatok ritkulása). Ezért a denzitás a globális értékláncok egy fontos mutatója, ami képes megmutatni az évenkénti értéklánc növekedést vagy éppen csökkenést.

A következő ábrán az ICIO adathalmazból képzett hálózat évenkénti denzitás értéke látszik 1995-től 2020-ig, ami szintén a globalizáció hatását igazolja, és az egyre összekapcsolt gazdasági hálózatok létrejöttét.



ábra 7: Denzitás értékek 1995-től 2020-ig, a teljes hálózatra

Ahogy az ábra mutatja világgazdaság denzitás értéke 1995-től folyamatosan emelkedett, azonban ez nem egy lineáris trendet követett. Néhány érdekes mintázat (időrendben): 2001-től a denzitás érték erőteljes növekedésnek indult, ami részben a kereskedelmi liberalizációnak köszönhető. Kína ebben az évben csatlakozott a WTO-hoz (World Trade

Organization), ez lehetővé tette a vámok csökkentését, a piaci elérés növelését, és még kiterjedtebb globális ellátási láncok létrehozását, így Kína a gyártás és a kereskedelem egyik központi csomópontjává vált, ezzel jelentősen növelve export és import tevékenységeit. Továbbá fontos szerepet játszott az Európai Unió bővülése 2004-ben, ami szintén a vámok jelentős csökkenését vagy adott esetben eltörlését is jelentette a csatlakozó tagállamok (Ciprus, Csehország, Észtország, Magyarország, Lettország, Litvánia, Málta, Lengyelország, Szlovákia és Szlovénia) esetében.

A 2008-as világválság a nagy pénzintézetek összeomlása a hitelek befagyasztásához, a fogyasztói kereslet csökkenéséhez és ezáltal a kereskedelmi volumen meredek visszaeséséhez vezetett. Az export orientált gazdaságok megszenvedték a globális kereslet visszaesését, míg a nyersanyagárak zuhanása az erőforrásoktól függő nemzeteket sújtotta. Továbbá a protekcionista politikák és a megzavart kereskedelem finanszírozás tovább gyengítette a nemzetközi kereskedelmet, amely 2009-ben több mint 12%-kal zsugorodott (WTO)

2014-ben az olajárak drasztikus zuhanásának hatására egy hullámvölgy keletkezett amit az ábra is mutat. Ennek az okai között az amerikai termelés miatti túlkínálat, valamint a gyengülő globális kereslet, különösen Kína részéről. Az árcsökkenés súlyosan érintette az olajexportőr országokat, például Oroszországot, Venezuelát és Nigériát, recessziót okozva ezekben a régiókban, és lassítva a globális gazdasági növekedést. Ezt azonban tovább súlyosbította a 2014-es Krím krízis és a Nyugat által Oroszországra kivetett szankciók, amelyek jelentős gazdasági zavarokat idéztek elő Kelet-Európában, amelynek negatív hatásai a szomszédos gazdaságokra is kiterjedtek.

A 2019–2020-as recessziós időszak, amelyet elsősorban a COVID-19 világjárvány okozott, fennakadásokat okozott a globális kereskedelmi hálózatokban. A lezárások, az utazási korlátozások és a gyárbezárások megzavarták az ellátási láncokat és csökkentették a termelési kapacitást. A globális kereslet zuhant, különösen a nem alapvető áruk iránt, miközben a szállítási késések és a növekvő kereskedelmi akadályok tovább nehezítették a kereskedelmi áramlásokat. A globális kereskedelem volumene 5,3%-kal esett vissza 2020-ban (WTO).

Tehát elmondható, hogy a gráf általában folyamatosan növekvő nem lineáris trendet mutat. Azonban ezt az imént ismertetett recessziók árnyalják, amik megmutatják a globális értékláncok sérülékenységét.

A következő elemzésekben a hálózathól egy többrétegű hálózatot készítettem, amin a rétegen belüli és a rétegek közötti kapcsolatokat vizsgáltam. A hálózatot országok, majd szektorok szerint alakítottam egy-egy multi-layer hálózattá.

3.6.1. Rétegen belüli denzitás (Intra-layer density)

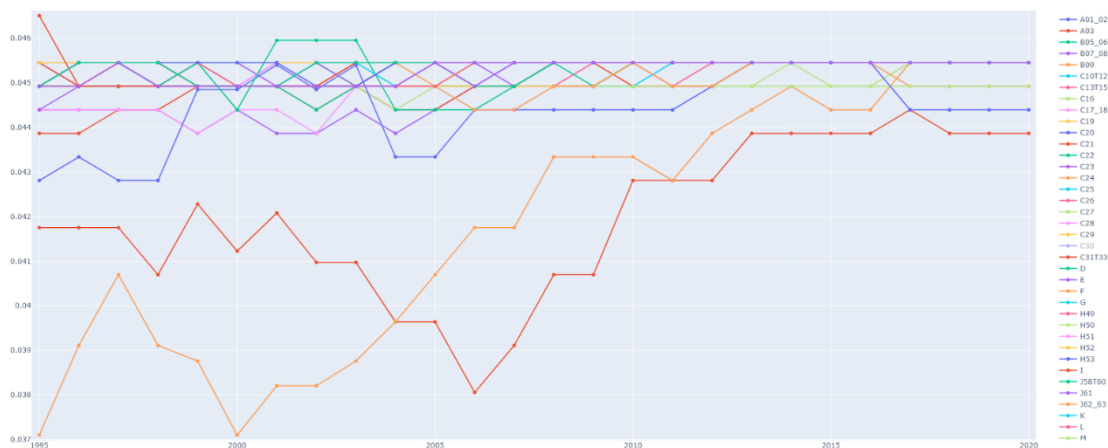
A rétegen belüli denzitás azt mutatja meg, hogy adott rétegen belül mennyire összekapcsoltak az adott csomópontok. Az első esetben, mikor a többrétegű hálózat rétegeit sz országok képezik, az országok szektorainak összekapcsoltságát mutatja a denzitás. Ezt idősoros diagramon vizualizálva érdekes mintázatok figyelhetők meg, amit a következő ábrákon ismertetek.



ábra 8: Top 10 ország rétegbeli denzitása 1995-től 2020-ig

Az ábrán az általam 10 legérdekesebbnek ítélt ország látszik, amik Kanada, Kína, Németország, Franciaország, az Egyesült Királyság, Indonézia, India, Olaszország, Japán és Korea voltak. Az országokon belüli intenzív kereskedelemnek (amit a magas denzitás jelez) köszönhetően a szektorok kevésbé lehetnek érzékenyek az embargók negatív hatásaira.

A diagramon ábrázolt időintervallum első harmadában például India gazdasága a belső piacra koncentrált, míg a következő időszakban Kanada főleg a külső piacaira fókuszált.



ábra 9: Top 10 szektor rétegbeli denzitása 1995-től 2020-ig

A szektorok szerinti bontás esetén, mivel a szektorok összekapcsoltsága nagyon hasonló volt így csak a 10 millió dollárnál magasabb értékű kereskedelmi utakat vettem figyelembe. Az így kapott ábrán már meg tudtam állapítani néhány érdekes motívumot (alsó 3 görbe, sorrendben):

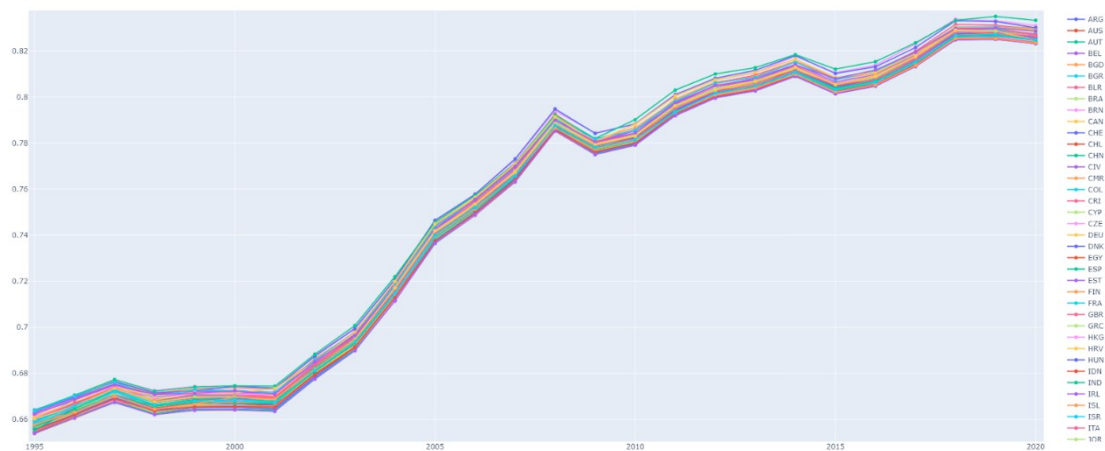
- [A03]Halászat és agrikultúra - Bár nem jelentősen, de a vizsgált időszakban 2006-ban mélypontot ért el a halászati szektorok összekapcsoltsága más szektorokkal, illetve ennek a szektornak a különlegessége, hogy általában a legalacsonyabb érték. Ez valószínűleg azért van, mert a mezőgazdasági és halászati javakat kevés és direkt szektorok dolgozzák fel.
- [B09]Bányászati segédszolgáltatás - Az ábrán egy 1995 óta növekvő tendencia látszik, ami azt jelzi, hogy a bányászati segéd szolgáltatásokat más szektorok is igénybe vették az idő múlásával.
- [H53]Posta és futárszolgálat - A vizsgált időszakban a postai szolgáltatások igénybevétele ingadozó volt, ez azt jelenti, hogy voltak olyan időszak amikor sok szektor vette igénybe a szolgáltatást (pl 2013-2016) és voltak olyan évek (pl 1995-1998) amikor kevesebben igényelték.

A többi szektor a vizsgált időszakban csak minimálisan változott, a denzitás értékeik végig magas volt.

Az eredmények azt mutatják, hogy az országok belső kereskedelme jelentős hatással van a szektorok stabilitására, különösen válságos időszakokban. A vizsgált időszakban érdekes változások voltak megfigyelhetők az egyes szektorokban.

3.6.2. Rétegek közti denzitás (Inter-layer density)

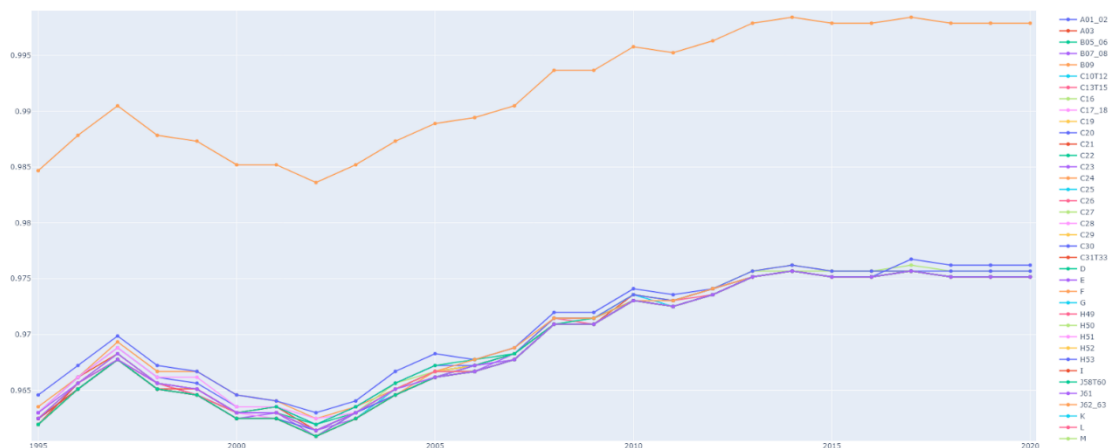
A rétegek közti denzitás azt mutatja meg, hogy a rétegek közötti (esetünkben a rétegek országokat vagy szektorokat jelentenek) kereskedelem mértéke mekkora, a teljes összekapcsoltsághoz viszonyítva. A következő ábrán az x tengelyen az idősor, míg y tengelyen a denzitás értékek láthatóak. Mivel esetünkben rétegek közötti kapcsolatokat vizsgálunk az ország itt exportáló országot jelent. Ebből adódóan a denzitásérték itt minden lehetséges export úthoz viszonyítva a meglévő kereskedelmi utak arányát mutatja meg.



ábra 10: Rétegek közti denzitás országonként 1995-től 2020-ig

Összességében az országok tekintetében elmondható, hogy az országok rétegek közti kereskedelme nagyon hasonló ütemben változott. A kapott gráf hasonló az egész hálózatra számított denzitásra, valószínűleg a gráf karakterisztikája az interlayer kapcsolatoknak köszönheti alakját.

Továbbá megvizsgáltam szektoronkénti bontásban is a rétegek közti kereskedelmi kapcsolatokat. Ahol egy outlier-t kivéve szintén hasonló volt a szektorok tendenciája.

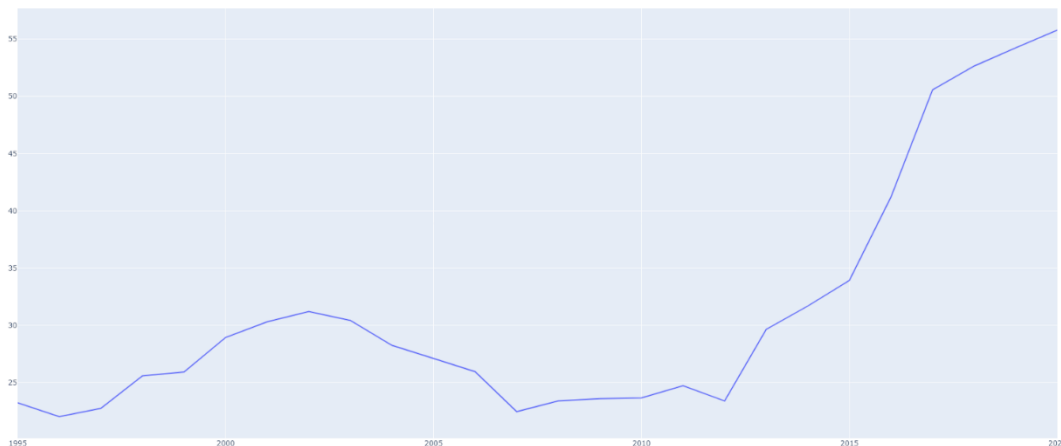


ábra 11: Rétegeközi denzitás szektoronként 1995-től 2020-ig

Az egy outlier szektor, a [T]Háztartások munkaadói tevékenysége, saját felhasználásra volt, ami meglepő.

3.7. Homofília (Homophily)

A homofília többrétegű kereskedelmi hálózatok esetén azt írja le, hogy a hasonló tulajdonságokkal rendelkező szereplők (országok vagy szektorok) nagyobb valószínűséggel kötnek egymással kereskedelmi kapcsolatokat, mint a különböző tulajdonságokkal rendelkező társaikkal.



ábra 12: Homofília változása 1995-től 2020-ig

Az ábrán 2015-től egy ugrásszerű növekedés látszik, ami mögött több geopolitikai esemény is húzódhat, a Krím annektálása (2014) ami a regionális gazdasági kapcsolatok megerősödéséhez vezetett, ezáltal egy blokkosodást eredményezett amit az imént említett homofília mutatóval lehet leírni.

Szektoriális téren is megfigyelhető egy blokkosodás, ha csak egy technológiai példát nézünk is megmutatkozik. Míg Amerikában és Európában elsődlegesen például Facebook-ot használunk közösségi média elérésre, Kínában a WeChat, Ruszin-szférában pedig a VKontatje-t elterjedt alternatíva.

A következőkben a kiszámolt központi mutatók segítségével az ICIO hálózatokon közösség-észlelési technikák alkalmazási lehetőségeit vizsgálom meg.

4. Közösségészlelés (Community Detection)

A közösségészlelés során az algoritmusok átfedő (overlapping) és nem átfedő (non-overlapping) módszerekre oszthatók, attól függően, hogy egy csomópont (ország/iparág) tartozhat-e több közösséghez.

4.1. Kereskedelmi közösségek azonosítása

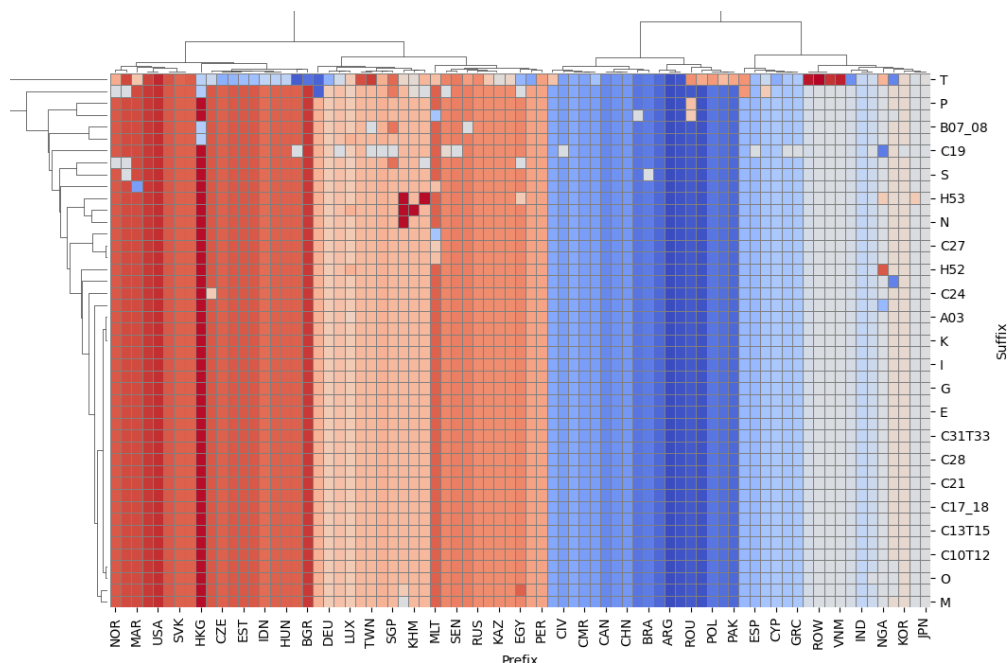
A kereskedelmi közösségek azonosítása segít feltárni a szoros gazdasági együttműködések és a regionális integrációkat a nemzetközi kereskedelmi hálózatokban. Az ilyen közösségek meghatározása lehetővé teszi a piaci trendek előrejelzését és a gazdaságpolitikai döntéshozatal támogatását. Az erős integrációval rendelkező országok stabilabbak lehetnek a kereskedelmi válságokkal szemben, míg a perifériára szoruló gazdaságok nagyobb kockázatnak vannak kitéve (Newman, 2006). Az ilyen elemzések révén pontosabb képet kaphatunk a globális gazdasági hálózatok működéséről, elősegítve a kereskedelmi megállapodások kialakítását (Fortunato, 2010).

4.1.1. Louvain modularitás

A Louvain modularitás kereskedelmi hálózatokon való alkalmazása alkalmas kereskedelmi blokkok és szoros kereskedelmi megegyezések detektálására, illetve ezáltal rangsorolására is.

Ezen túlmenően, a Louvain modularitás segítségével összehasonlíthatóak a különböző időszakok, így kimutathatók a globalizáció, protekcionizmus vagy gazdasági válságok együttműködésekre gyakorolt hatásai. Az algoritmus alkalmazása tehát nemcsak a

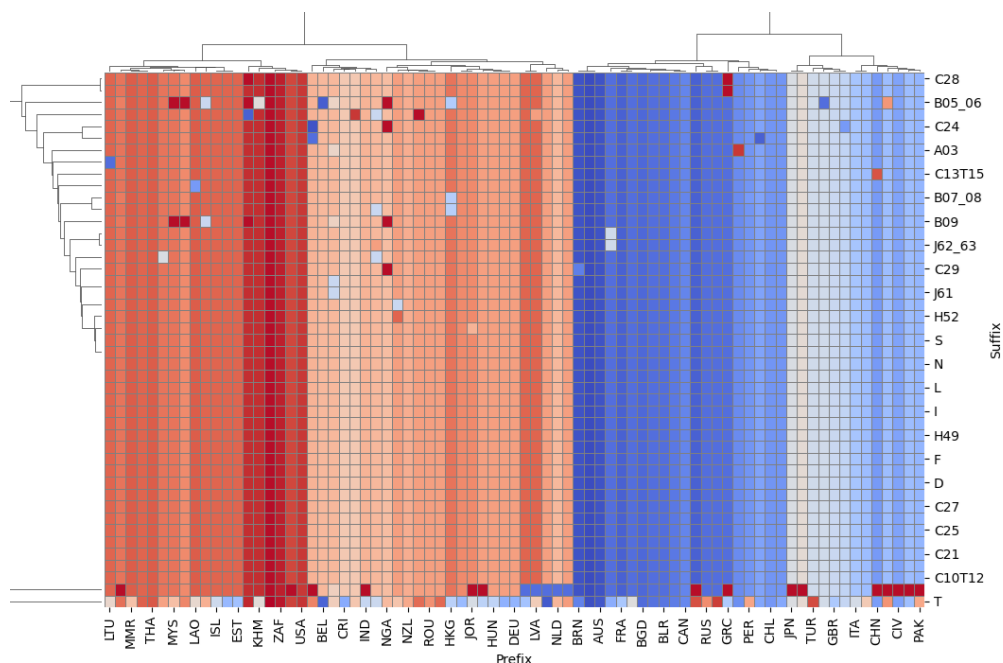
statikus struktúrák feltérképezésére alkalmas, hanem dinamikus változások nyomon követésére is.



ábra 13: Louvain algoritmus által megállapított modulok 1995-ben

A Louvain algoritmus 128 csoportot állapított meg 1995-ben, azonban ebből, csak 37 csoport rendelkezett egy elemnél több taggal. A különböző csoportokat az ábrán eltérő szín jelöli, ami esetünkben 128 színt jelent.

Ezek a megállapított csoportok, ahogy az ábrán is látszik általában egy ország szektorai voltak, ami nem meglepő. Azonban két szektor esetén ez nem állt fent, ami a [T]Háztartások munkaadói tevékenysége, saját felhasználásra és a [C19]Koksz és finomított kőolajtermékek volt. Ezek a szektorok jellemzően kissé szektor szintű csoportosulást követnek, mint ország szintűt, a többivel ellentétben.



ábra 14: Louvain algoritmus által megállapított modulok 2020-ban

2020-ra a Louvain algoritmus 127 csoportot állapított meg, ebből 34 csoport rendelkezett egy elemnél több taggal. Ebben az évben szintén jellemzően két szektorra volt jellemző az előző ábrán megállapított trend, ami a T és C19 szektor átrendeződése.

4.1.1.1. Ország-aggregált hálózat

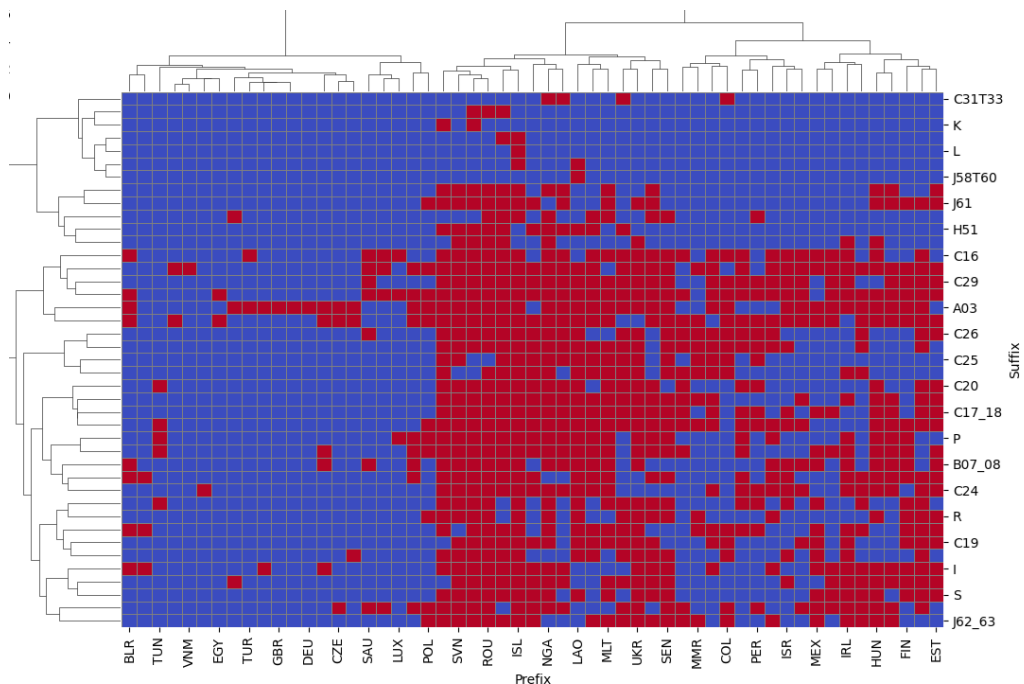
A multilayer hálózati elemzés esetén az aggregált hálózatot érdemes vizsgálni. Esetünkben az országok szerinti aggregáció esetén a Louvain algoritmus 5 csoportot alakított ki, ami az alábbiak szerint alakult 1995-ben.

- Szenegál, Ausztrália, Belgium, Bangladesh, Tunézia, Kínai Tajpej, Brazília, Kanada, Egyesült Államok, Brunei, Elefántcsontpart, Vietnám, Kolumbia, Costa Rica, Spanyolország, Egyesült Királyság, Görögország, Izland, Japán, Laosz, Málta, Nigéria, Hollandia
- Szingapúr, Argentína, Írország, Thaiföld, Luxemburg, Marokkó, Mexikó, Malajzia, Norvégia, Pakisztán, Fülöp-szigetek, Hongkong
- Szlovákia, Németország, Oroszország, Törökország
- Portugália, Szaúd-Arábia
- Kambodzsa, Korea, Chile, Peru
- Ausztria, Svédország, Szlovénia, Bulgária, Fehéroroszország, Ukrajna, Svájc, Dél-Afrika, Ciprus, Csehország, Dánia, Egyiptom, Észtország, Finnország,

Horvátország, Magyarország, Indonézia, Izrael, Olaszország, Jordánia, Litvánia, Lettország, Új-Zéland, Lengyelország, Románia

4.1.2. Mag-periféria struktúrák

A mag-periféria struktúra (core-periphery structures) egy kereskedelmi hálózatban megmutatja a gazdasági kapcsolatok hierarchiáját, ahol a “mag” erősen összekapcsolt országokból és szektorokból áll, míg a periféria kevésbé kapcsolódó, függő szereplőket tartalmaz. A mag országai/szektorai dominálják a globális kereskedelmet és formálják a globális ellátási láncokat, míg a periféria többnyire nyersanyagok és alacsony hozzáadott értékű termékek exportjára épül. Ennek az elemzésnek az előnye, hogy feltárja a kereskedelmi függőségi viszonyokat, a kereskedelmi hálózat stabilitását, valamint ha időrendet veszünk figyelembe a sokkterjedési mechanizmusokat és képes megmutatni. Hátránya viszont, hogy a modell merev hierarchikus struktúrát feltételez, nem mindig tükrözi a dinamikus gazdasági változásokat, és kevésbé figyel a regionális együttműködésekre.



ábra 15: Mag-periféria által megállapított mag-határok 1998-ban

Az ábrán biklaszterezést is alkalmaztam, így általánosságban a hőterkép jobb alsó harmadába kerültek a mag, ezen kívül pedig a periféria országok/szektorok. Ezáltal könnyen megfigyelhetőek a mag országok (és szektorok) amik központi szerepet töltenek

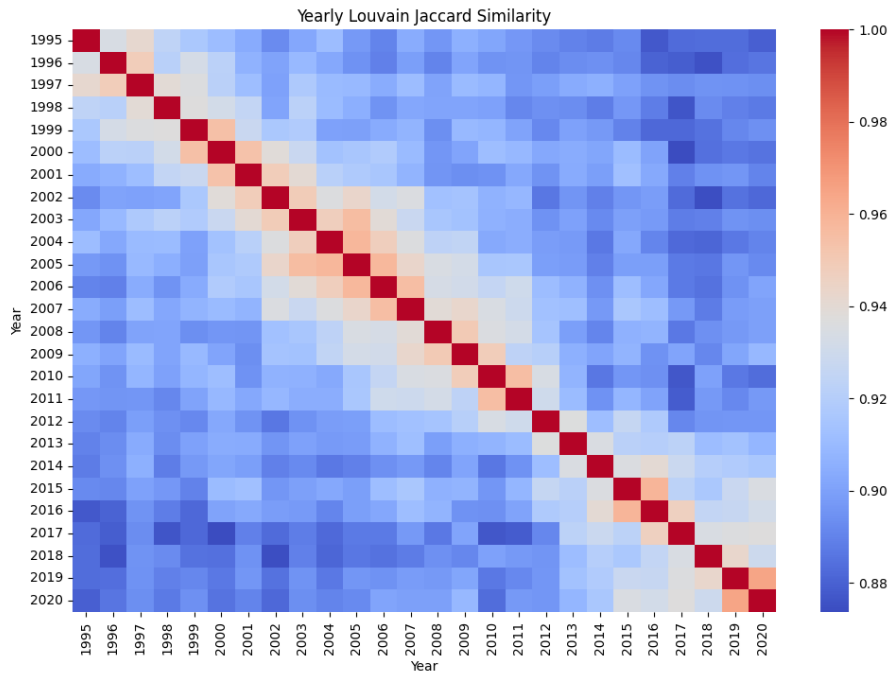
be a világkereskedelemben, illetve a periféria országok (és szektorok) amik kevésbé jelentős szerepet töltenek be.

4.2. Közösségek idősor analízise

A közösségek idősor analízise során az első pontban azonosított közösségeket időbeli változásuk szerint vizsgáltam. Az adathalmazon elemezhetjük, hogyan változnak a gazdasági együttműködések a kijelölt évek során, beleértve a globalizáció és a regionális szövetségek alakulását, ennek hatásait. Vizsgálhatjuk, mely országok és iparágak gazdasági teljesítménye mutat szorosabb együttmozgást az idő múlásával, továbbá megállapíthatjuk, hogy a gazdasági közösségek mennyire stabilak vagy éppen folyamatosan változó struktúrák jellemzik-e őket, és hogy milyen hatással vannak rájuk a válságok, technológiai fejlődés vagy kereskedelmi konfliktusok.

4.2.1. Dinamikus közösség detektálás (Dynamic Community Detection)

A dinamikus közösség detektálás arra szolgál, hogy az időbeli változásokat is figyelembe tudjuk venni. Általa megvizsgálhatjuk, mennyire állandóak az egyes gazdasági közösségek, azonosíthatjuk, hogy bizonyos események (válságok, világjárványok, geopolitikai változások) milyen hatással vannak a gazdasági közösségek szerkezetére, illetve feltárhatjuk, hogy léteztek-e átmeneti időszakok, amikor a gazdasági blokkok átalakultak, szétestek vagy újjászerveződtek. Mivel a Louvain algoritmus egy statikus közösség detektáló eljárás, így erre önmagában nem alkalmas, azonban ha idősorra alkalmazzuk majd a számított Louvain-idősorra valamilyen hasonlóságot számítunk, megkaphatunk az évek között egy hasonlósági metrikát. A következőkben Jaccard távolságot számítottam, amit “best-match” azaz legjobb párosításban alkalmaztam (két év között a két leghasonlóbb közösséget párosította). Az így kapott ábra a következőképp alakult:

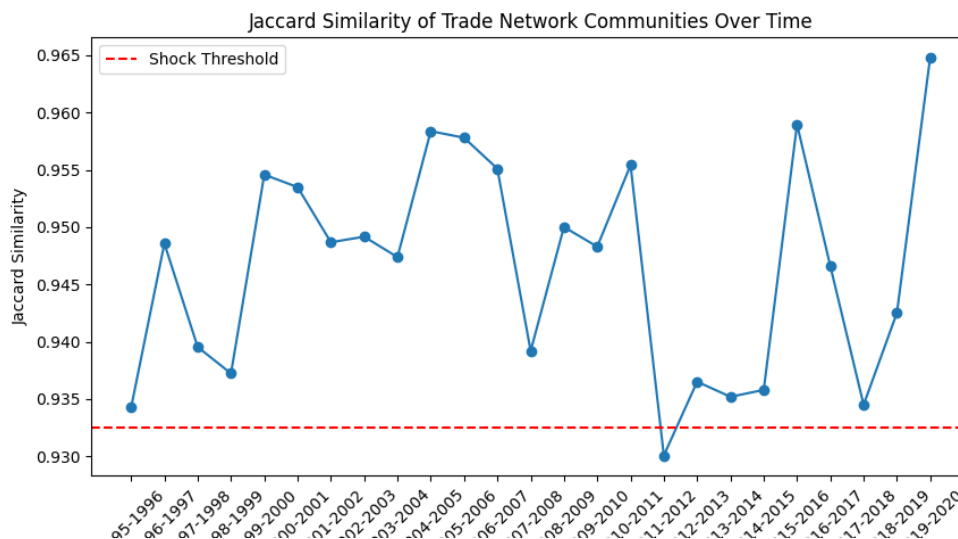


ábra 16: Jaccard távolság az évekre számított Louvain modularitásra

Az egymás után következő évek általában jobban hasonlítanak egymásra mint a távolabb lévő évek, tehát elmondható, hogy kis időtávon általában kis átalakulások voltak jellemzők, azonban nagy időtávon már jellemzően kevésbé voltak jellemzőek hasonló csoportok. Egy érdekes jelenség, hogy a legnagyobb átrendeződések 2016-17-ben történtek az előző időszakokhoz képest, amit magyarázhat többek között Donald Trump beiktatása, amivel az Egyesült Államok egy protekcionista fordulatot vett, a Kínai "Egy övezet, egy út" (Belt and Road Initiative) vagy akár a Brexit-tel.

4.3. Anomália-észlelés

A sokkelemzés célja hogy azonosítsa és elemezze a hirtelen változásokat egy hálózatban. Esetünkben az előzőekben számított Louvain-Jaccard közösség távolságokat felhasználva vizsgáltam meg az egymást követő év-párokat, amit a következő ábra szemléltet.



ábra 17: Sokkelemzés a DCD-nél azonosított értékekre

A sokkelemzés során a küszöbértéket a következő képlettel állapítottam meg:

$$(threshold = \mu - 1.5 \cdot \sigma)$$

(1)

Ahol az átlag (μ) az összes év-pár Jaccard-értékeinek átlaga, a szórás (σ) pedig a Jaccard távolságok “változékonyságát” méri. A szórást megszoroztam 1,5-tel, ami egy lazább kritérium mint az általában használt 2-es érték (amit általában anomália-detektálásnál használnak), ez elég szigorú ahhoz, hogy csak nagyobb kereskedelmi átalakulásokat jelezzon, de még nem túl szigorú, hogy semmit se detektáljunk. Így egy fordulóra állapított meg sokkot az algoritmus, ami a 2010-2011-es évek voltak. Ami magyarázható az arab tavasszal, ami során az arab országok külkereskedelme növekedett, vagy akár a 2008-as gazdasági világválságból való kilábalással.

4.3.1. Autoencoder-alapú anomália-észlelés

Az Autoencoder-alapú anomália-észlelés egy olyan technika amit számos más területen használnak, mint például a kiberbiztonság, az ipari karbantartás és az egészségügyi diagnosztika. Ennek a módszernek az alapja, hogy az adatok általános mintázatát tanulja meg a hálózat, így ha egy adat jelentősen eltér a többitől a visszaalakítási hibája megnő. Kereskedelmi hálózatokon való alkalmazásához azonban némi átalakítást kell végezni, hogy a bemeneti réteg megfelelően tudja kezelni az adatot. Ezért az évekből képzett kereskedelmi hálózatoknak csak a hálózatok meghatározó jellemzői kerülnek felhasználásra és ezek képzik a neurális háló be- és kimentetét.

Az Autoencoderek hatékonyan megtanulják a bonyolult, nem lineáris összefüggéseket is. Emiatt jól alkalmazható magas dimenziójú, komplex adatszerkezetek vizsgálatára, mint amilyen az általunk használt ICIO adathalmaz is.

4.3.1.1. Algoritmus főbb lépései

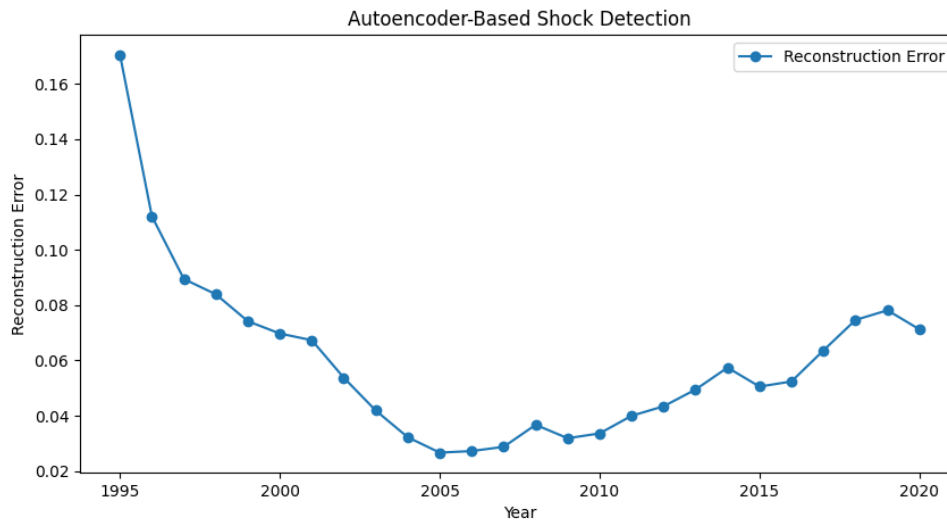
Az algoritmus megvalósításának az első része (az évenkénti kereskedelmi hálózatok betöltése után) a főbb hálózati jellemzők kinyerése, amit többek között a dolgozat első felében kiszámolt hálózati mutatók képeznek, apróbb kiegészítéssel. Ezek a mutatók a Fokszám, Súlyozott fokszám, Klaszterezési együttható, Közöttiségi központiség, PageRank, Sajátvektor központiség, K-mag szám. A felsorolt változók a sokkok általános detektálásáért felelősek, kivéve a PageRank, Sajátvektor központiség, illetve a klaszterezésért felelős mutatók. A PageRank és a Sajátvektor központiség a kulcsszereplők változásának a detektálásával bővíti az autoencodert, a klaszterezési mutatók (Klaszterezési együttható és a K-mag szám) pedig a regionális mintázatok változásával.

Az így kapott jellemző-mátrixot a kiegyensúlyozás miatt (hogy ne domináljon egyik változó sem) normalizáltam. Az így keletkezett normált adatok képezik az autoencoder bementét.

4.3.1.2. Autoencoder felépítése

Az autoencoder tradicionális encoder-decoder struktúrát alkalmaz amit PyTorch segítségével valósítottam meg. A meghatározott jellemző-mátrix alapján egy 7 neuronnal rendelkező réteg képezi a hálózat bemenetét, amit egy rejtett réteggel 3 dimenzióssá alakított az encoder. Amit aztán egy decoder ennek egy fordított struktúrájával megpróbálja a 3 dimenziós "tömörített" adatot visszaállítani az eredeti állapotába. A hálózat tanítása során ADAM optimalizációt használtam 0.01-es tanulási rátával, ami az átlagos négyzetes hiba (MSE) minimalizálásával tanítja a hálózatot. A hálózat kiértékelése során az eredeti jellemző-mátrixot évenként összehasonlítottam az autoencoderrel beágyazott majd visszaállított mátrix-szal. Ez az eltérés a rekonstrukciós hiba, ami az outlier éveket jelöli és amik a feltételezett sokkok lehetnek. Az így kapott eredményeket egy vonal-diagramon ábrázoltam.

4.3.1.3. Eredmények kiértékelése

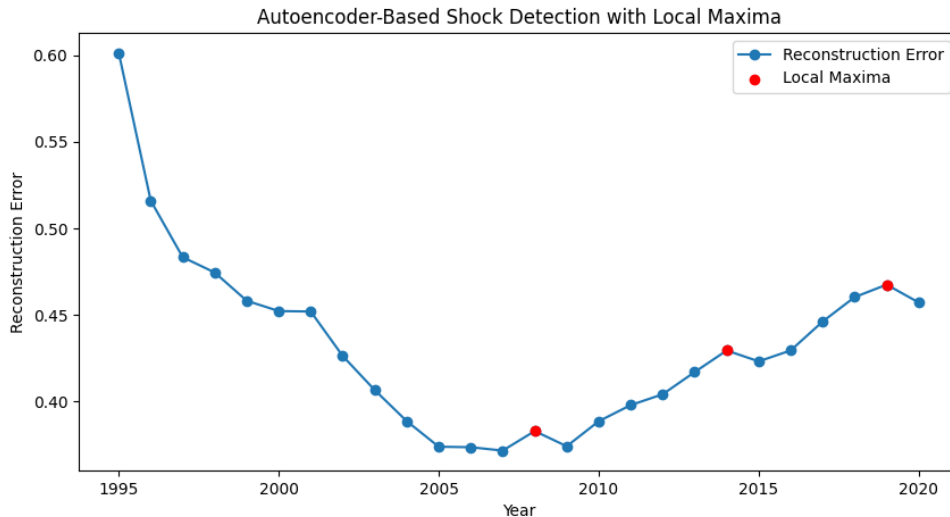


ábra 18: Autoencoder által megállapított sokkok 1995-től 2020-ig

Az ábrán megfigyelhető kezdeti kiemelkedő rekonstrukció hiba feltételezhetően abból is adódhat, hogy a kereskedelmi hálózat 1995-96-ban sokkal kisebb volumenű nemzetközi kereskedelmet bonyolított le, mint az azokat követő években.

4.3.1.4. Lokális maximumkeresés

A grafikonon megállapítható V alakú görbület valószínűleg abból adódik, hogy az autoencoder által használt központosági mutatók nem homogén változásokat tartalmaznak. Ez azt jelenti, hogy egyes mutatókban az idősorban a gazdasági növekedés erőteljesebben jelentkezik, míg másokban kevésbé. Ennek a kiküszöbölésére és a rövid (1-2 év) sokkok detektálására megoldást jelenthet a lokális maximum keresés. Az elemzésem során a scipy argrelextrema függvényét használtam, ami egy páros összehasonlító algoritmus, esetünkben egy adott évet hasonlít össze a közvetlen szomszédaival.



ábra 19: Lokális maximák a Sokkelemzés során megállapított gráfon

A maximum-keresés eredményeként 2008, 2014 és 2019-es éveket jelölte meg az autoencoder alapú megoldás.

- 2008: A pénzügyi világválság miatt a globális kereskedelmi forgalom drámai csökkenésen ment keresztül.
- 2014: Az olajár-esés mély gazdasági hatásokat váltott ki, különösen az olajexportáló országokban.
- 2019: Amerikai-Kínai vámháború és az ebből következő gazdasági bizonytalanság miatt a világkereskedelem növekedése lelassult.

Egy további érdekes jelenség, hogy az ábra első felében, a 2000-es évek elején a 2001-es év ami emberi szemmel maximának tűnik (de nem az) valószínűleg a .com buborék (dot-com bubble). Ennek a csúcsnak a nem felismerése származhat abból, hogy a .com buborék egy hosszabb folyamat volt, nem egy hirtelen sokk, illetve az imént említett trend is zavarta a felismerést. A hálózat tanítása és az eredmények kiértékelés nagyban függhet a hálózat inicializálásától, de általában az imént említett csúcsokat detektálta a hálózat sokként.

4.4. kereskedelmi blokkváltások előrejelzése

A következőkben a kereskedelmi blokkok szerkezeteinek idősoros változását vizsgálom, amivel modellezhetőek az előző módszerek által fel nem ismert változások.

Erre egy a már az autoencoder alapú anomália-észlelés során használt technikát, a beágyazást (embedding) fogom használni. A csomópont-beágyazás (node embedding)

alkalmazása a kereskedelmi hálózatokban lehetővé teszi az országok közötti komplex kereskedelmi kapcsolatok numerikus vektortérben történő reprezentálását. Ez a megközelítés megőrzi a hálózat struktúráját és a csomópontok közötti kapcsolati mintázatokat, ezáltal a kereskedelmi blokkok reprezentálása (és azok változásai) is lehetséges egy kisebb dimenziójú vektortérben.

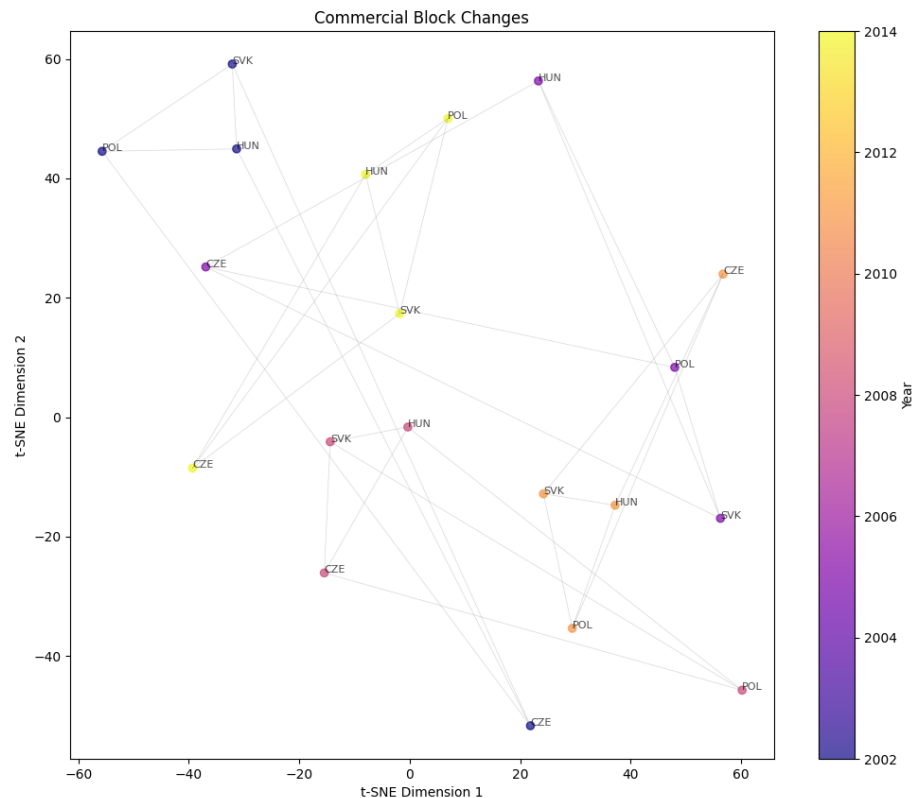
4.4.1. Node2Vec és t-SNE

A hálózatok beágyazását a node2vec függvénnyel valósítottam meg, amivel a csomópontokat egy magasabb dimenziójú vektortérben helyeztem el (az ábrák készítésekor 64 dimenziót alkalmaztam). Az algoritmus a véletlen bolyongás elvén működik, ami az jelenti, hogy a gráfban véletlen utakat keres és az így keletkezett szekvenciákat dolgozza fel. A bolyongást befolyásoló paraméterek, többek között, a megteendő utak száma, hossza, illetve a Q és P paraméterek. Ha a Q paraméter magas, a bolyongás inkább a lokális struktúrához ragaszkodik, ha alacsony, akkor inkább globálisan fedezi fel a gráfot. A P paraméter pedig a visszatérési paraméter, ami ha magas, az algoritmus kevésbé hajlamos visszatérni az előző, már meglátogatott csomóponthoz. Az így keletkezett csomópont-szekvencia halmaza, pedig majd a skip-gram modellel egy alacsonyabb vektortérbe vetíti. A skip-gram a szekvenciák alapján megtanulja, hogy egy adott csomópont milyen kontextusban (mely más csomópontok közelében) fordul elő. Az így kapott 64 dimenziós vektortér az ábrázolhatóság miatt egy 2D térbe vetítettem a t-SNE (t-Distributed Stochastic Neighbor Embedding) algoritmus segítségével.

A t-SNE egy dimenziócsökkentő algoritmus, amely különösen hasznos magas dimenziós vektorok vizualizálására alacsonyabb dimenziós térben. Az algoritmus a távolság-rangok megőrzésére törekszik, így alacsonyabb dimenziós térben is megtartja az adatok helyi szerkezetét. Az algoritmus először Gauss-eloszlásokkal kiszámítja az adatpontok közötti (páronkénti) hasonlóságokat a magas dimenziós térben, majd egy Student-féle t-eloszlással (t-SNE) modellezi ezeket az alacsony dimenziós térben. A két tér közötti különbség minimalizálása KL-divergenciával történik, ami egy valószínűségi eloszlások közti különbséget mérő függvény, a minimalizálásával a közeli pontok hasonló valószínűséggel maradnak közel egymáshoz, míg a távoli pontok elkülönülnek.

4.4.2. Eredmények kiértékelése

Mivel a beágyazás során a kereskedelmi hálózat tulajdonságai megőrződnek a csomópontok távolságaként, így ez a módszer könnyen használható a csomópontok közötti hasonlóságok észlelésére, ami ezáltal lehetővé teszi a kereskedelmi blokkváltások előrejelzését. Ez fontos a kereskedelmi hálózatok dinamikus természetének megértésében és a stratégiai döntéshozatal támogatásában.



ábra 20: Beágyazott eredmények V4-es országokra 2002-től 2014-ig (3 évente)

Az ábrán látható V4-es (Visegrádi Négyek) országok Csehország, Lengyelország, Magyarország és Szlovákia 2002-től 2014-ig az idő múlásával egyre szorosabb köteléket alkottak a beágyazás szerint. Az évek során azonban látható, hogy Csehország általában távolabb fekszik a V4-es partnereitől, ami azért lehet, mert gazdasága erősebben kötődik Németországhoz és Nyugat-Európához, míg a többi V4-ország inkább egymással és szomszédaikkal kereskednek. Történelmileg is szorosabban kapcsolódik a német ipari ellátási láncokhoz, különösen az autóipar terén.

5. Gráf Neurális Hálók alkalmazhatósága

A gráf neurális hálózatok (GNN - Graph Neural Network) különlegessége a konvencionális neurális hálókhoz képest leginkább az input és output rétegekben jelentkezik. A GNN-ek bemenetét ugyanis a hagyományos vektorokkal ellentétben magák a gráf objektumok, a használt PyTorch Geometric (PyG) esetében PyGData típusú gráf reprezentációs objektumok.

Az általam készített neurális hálók konvolúciós rétegeket alkalmaznak amivel úgy készítenek alacsonyabb (encoder) vagy magasabb (decoder) dimenziós reprezentációkat, hogy figyelembe veszik a csomópontok szomszédsági szerkezetét.

A GNN-ek mint koncepció már a 2000-es évek előtt is léteztek azonban csak az utóbbi években lettek igazán népszerűek hála a gráf konvolúciós (GCN - Graph Convolutional Networks) és a gráf figyelmi (GAT - Graph Attention Networks) hálózatoknak.

5.1. Gráf Neurális Hálózatok alkalmazása kereskedelmi hálózatokra

A dolgozat eddigi részeiben olyan megoldásokat mutattam be amik beágyazzák és visszaalakítanak gráfokat, azonban ezek csak központi mutatókkal voltak képesek (autoencoder), illetve előre kidolgozott vektortérbe (node2vec) voltak képesek beágyazni. A gráf neurális hálózatok segítségével azonban a kereskedelmi gráfok átalakítás nélkül átadhatók a hálózatnak, így minden tulajdonságára képesek rátanulni, még olyanokra is amik a használt központi mutatókkal nem mutathatók ki.

5.1.1. PygData

Az említett PyGData objektum a gráf egy speciális reprezentációját tárolja, aminek 3 fontos attribútuma van. Az **X**, ami a csomópontok jellemzői, esetünkben ország és szektor index (numerikus reprezentáció engedélyezett), az **edge_index** attribútum ami az élek kapcsolati listáját tárolja (például: ARG_A03 -> GER_C22) és az **edge_attr** ami az élekhez tartozó súlyokat tárolja, ami esetünkben a két ország két szektora között folyó kereskedelem mennyisége.

5.1.2. Konvolúciós rétegek

Az implementált hálózatok gráf konvolúciós hálózatok, ami mind a autoencoder, mind az embedding hálózat esetén 2 konvolúciós réteget tartalmaz, amivel a hálózatok még kiterjedtebb reprezentációt képesek megtanulni a különböző csomópontokra. Ugyanis a konvolúciós rétegek GNN viszonylatban azt jelentik, hogy minden csomópont a szomszédos csomópontokból és a saját jellemzőiből készít egy reprezentációt, ami egy új beágyazás ami már az előző lokális struktúrát tükrözi. Az új beágyazás elkészítésekor természetesen egy aggregációra is szükség van ami általában súlyozott átlagolás vagy valami egyéb transzformáció.

5.2. Kereskedelmi hálózatok beágyazása GNN segítségével

Az általam kidolgozott megoldás egy olyan GNN-t alkalmazó rendszer, amely az összetett ICIO kereskedelmi hálózat beágyazásán keresztül segít megérteni a gazdasági dinamikákat. A hálózat a tanítási folyamat során egy vektorteret képez, amiben később újra elhelyezi a vizualizálni kívánt választott csomópontokat.

5.2.1. Algoritmus főbb lépései

A függvény első részében a kereskedelmi gráfok átalakítás történik NetworkX objektumról a már említett PyGData objektummá. Majd az így kapott gráf-reprezentációkat a definiált beágyazási hálózatnak adjuk tovább tanulásra. A tanulási folyamat az eredeti autoencoder megoldáshoz hasonlóan ADAM optimalizálót használ MSE visszaterjesztéssel. Majd a második lépésben a feltanított modellre ráillesztjük a kívánt ország és szektor csomópontokat a kívánt évek szerint. Az így kapott magas dimenzionalitású (esetünkben 16) adatokat pedig a szintén a korábban elkészített autoencoderhez hasonlóan T-SNE segítségével a távolságrankok megtartásával 2 dimenziós térben vizualizáltam.

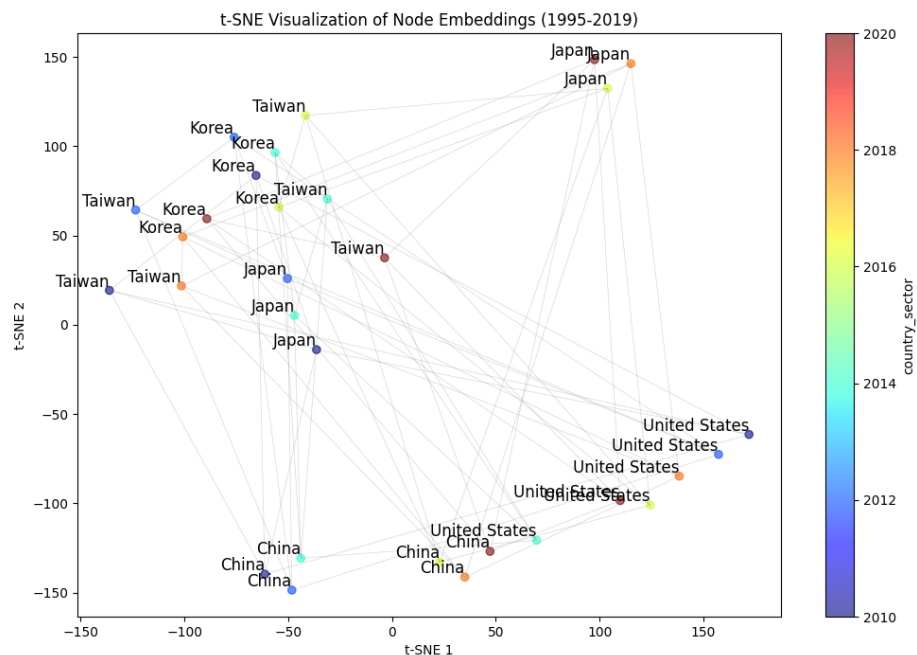
5.2.2. A hálózat felépítése

A hálózat egy a tanulási folyamat során egy szektor szerinti és egy ország szerinti beágyazást tanul meg, amit később összegez, így képes a látens reprezentáció

létrehozására. A hálózat két konvolúciós réteget tartalmaz, az első konvolúciós réteg az ország- és szektorbeágyazásokat egy közös rejtett reprezentációvá alakítja, míg a második lekicsinyíti ezt egy végső beágyazási dimenzióra. A modell, mivel figyelembe veszi a súlyokat (azaz a kereskedelem mennyiségét) így ez is befolyásoló tényező a beágyazás során.

5.2.3. Eredmények kiértékelése

A már említett t-SNE segítségével képzett dimenzió-csökkentett eredmények vizualizálásával a következő eredményeket kaptam. A diagrammon ha közel van két ország beágyazott csomópontja, az azt jelenti, hogy hasonló mintázattal kereskednek, vagy sűrűn, illetve intenzíven kereskedtek egymással.



ábra 21: Beágyazott csomópontok, az elektronikai ipar fontosabb szereplői

Az ábrán 2010-től 2020-ig két éves bontásban 5 ország látszik a 17-es szektor (számítógépes, elektronikus és optikai berendezések) tekintetében. Az 5 ország Kína, Egyesült Államok, Korea, Japán és Tajvan. Azért ezeket az országokat választottam az elektronikai szektor tekintetében, mert mindegyik ország jelentős szereplő például a félvezető gyártásban, de az elektronikai szektor más részeiben is jelentős szereplők. Az ábrán, habár véletlenszerű inicializációval történik a generálás általában megfigyelhető, hogy Korea, Japán és Tajvan szorosan együttműködnek az elektronikai szektorban, hiszen egymáshoz közel helyezkednek el. Ez tükrözi az ázsiai elektronikai ellátási lánc

szoros integrációját, ahol ezek az országok egymás iparába szállítanak be. Kína és az Egyesült Államok ugyan egy sarokban vannak, de inkább kétoldalú, kevésbé integrált kapcsolatuk van a többi ázsiai országhoz képest és valószínűleg a nagy volumenű kereskedelem miatt kerültek egymáshoz ilyen közel. Japán különösen érdekes, mert sokszor majdnem középen helyezkedik el, ami azt jelzi, hogy több ország felé is jelentős kereskedelmi kapcsolatokat tart fenn Tajvan és Kína között egy közvetítő szerepet is betölt. Japán politikailag egy semleges pozíciót vesz fel, mert habár elfogadja az egy-Kína (one-China-policy) politikát Tajvannal is szoros gazdasági kapcsolatokat ápol, az ábra pedig arra enged következtetni, hogy Tajvannal még szorosabb kapcsolatokat is ápol az elektronikai szektoraikban.

5.3. Közösségészlelés

Az előző részben készített beágyazó GNN modellre építve kibővítettem a függvényt, egy osztályozási résszel. A beágyazott gráfok osztályozásával és az osztályok évenkénti változásával jobban meg tudjuk érteni a nemzetközi kereskedelmi gráfok szerkezetét és azok időbeli változásait. Ez a megoldás azért működhet jól, mert a beágyazás során a csomópontok amik hasonló tulajdonságokkal rendelkeznek és szorosabb kereskedelmi kapcsolatokat ápolnak közelebb kerülnek egymáshoz. Az így nyert közösségek nemcsak strukturális hasonlóságot mutatnak, de sok esetben gazdaságföldrajzi vagy ágazati összefüggéseket is tükröznek, azonban ez az ICIO adathalmazon kevésbé volt releváns, mert a legfőbb tényező a kereskedelmi volumen volt.

A dimenzió-csökkentésre ez esetben azonban főkomponens elemzést (PCA - Principal Component Analysis) használtam. Az osztályozáshoz pedig K-Means algoritmust használtam, mert a klaszter formája és elhelyezkedése miatt ez adta a legjobb eredményeket. A DBSCAN a pontok sűrűsége miatt kevésbé adott jó eredményeket, ugyanis az ábrán ~3500 csúcs látható (77 ország * 45 szektor).

5.3.1. Motiváció

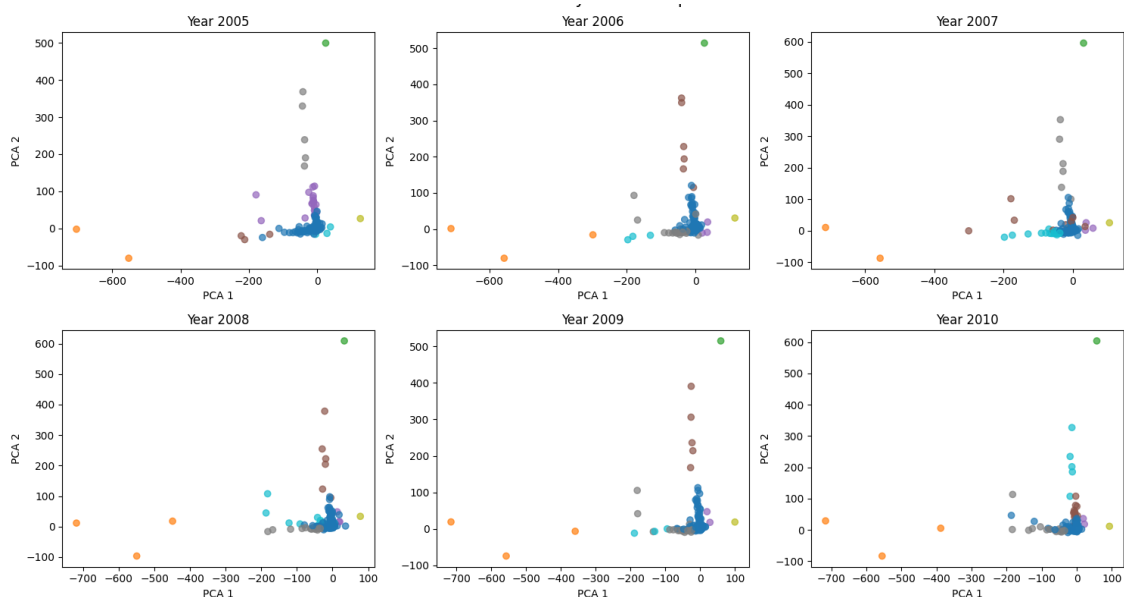
A GNN-alapú közösségészlelés hatékony lehet az ICIO és általánosságban a kereskedelmi gráfok esetében, mivel képes nemcsak a közvetlen kapcsolatokat és egyszerű hálózati központosági mutatók által megállapított metrikákat, hanem a magasabb szintű, többélű mintázatokat is képes figyelembe venni a reprezentációk során. A

magasabb szintű mintázatok (geometriai alakzatok) könnyebben is értelmezhetőek emberi szemmel, illetve ezek változásai is leírhatóbbá teszik a komplex világgazdasági eseményeket is. Ezek a magasabb szintű mintázatok, mint például az imént említett geometriai alakzatok, motívek vagy ismétlődő alstruktúrák fontosak nem csak az említett bilaterális kapcsolatok, hanem a komplex ellátási láncok leírása mentén is (könnyebb a mélyebb struktúrák feltérképezése is).

Bár ezzel a módszerrel előrejelzés nem lehetséges, utólagos kiértékelésre jól használható, például ha egy adott klaszterhez tartozó országok közötti kereskedelmi kapcsolat gyengül (távolabb kerülnek egymástól) vagy éppen új struktúrák alakulnak ki, az jelezhet politikai változást, új regionális együttműködést (beágyazás példában a V4-es országok) vagy akár gazdasági válságot is.

5.3.2. Eredmények kiértékelése

Mint említettem a közösségészlelés esetünkben idősoros elemzésben tud valóban hasznos eredményeket felmutatni így a következőkben elsősorban ezt vizsgáltam.



ábra 22: Klaszterek megállapítása K-means segítségével 2005-től 2009-ig

Az ábrán 2005-től 2010-ig az összes ország összes szektora látszik, azaz ~3500 csúcspont. A klaszterezés így inkább egy nagy csoportba sorolta az eredményeket.

1. Klaszterek idősoros táblázata 2005-től 2010-ig

Év	A Klaszter	B Klaszter	C Klaszter	D Klaszter	E Klaszter	F Klaszter	G Klaszter	H Klaszter
2005	3431	2	1	19	3	5	1	3
2006	3413	3	1	3	6	35	1	3
2007	3416	2	1	3	15	6	1	21
2008	3419	3	1	3	6	20	1	12
2009	3418	3	1	3	5	31	1	3
2010	3415	3	1	3	15	22	1	5

5.4. Anomália-észlelés

Egy másik érdekes felhasználása a gráf neurális hálózatoknak az anomália-észlelés, ez a korábbiakban említett autoencoder megoldáshoz képest egy sokkal jobb megoldás, ugyanis míg az előző autoencoder bemenetét csak az egész hálózatra megállapított központisági mutatók képezték a GNN alapú megoldás habár ennél komplexebb, pontosabb eredményeket tud adni. Ugyanis a GNN alapú hálózat bemenetét maga a gráf képezi így rejtett tulajdonságokat is figyelembe tud venni a hálózat ami a centralitási mutatókra történő “dimenziócsökkentés” esetén elvesztek.

Mivel a gazdasági sokkok kevésbé gyakran fordulnak elő, és mivel az autoencoder a “gyakori” struktúrákat tanulja meg, jól meg tudja különböztetni a zajt (potenciális sokk) és pont ezért tudja megállapítani azokat az éveket, amik nem illeszkednek a megszokott mintákba, amik általában gazdasági sokkra vagy hirtelen fellendülésre utal.

5.4.1. A hálózat felépítése

A már bemutatott módszerhez hasonlóan itt is a rekonstrukciós hibából ítélhető meg a gazdasági sokk. Azaz a modell saját magától tanulja meg, hogy mi számít „szokásosnak”. Az anomáliák ott bukkannak fel, ahol a rekonstrukciós hiba jelentős, vagyis a tanult minták nem tudják jól visszaadni az adott gráfot.

A készített autoencoder hálózat időfüggetlen, ami azt jelenti, hogy habár minden évhez külön rekonstruált gráf tartozik, ugyanarra a modellre végeztük a tanítást így egy közös beágyazási tér keletkezik. Ezzel könnyen összevethetők az évek között a beágyazások és még fontosabb, a rekonstrukciós hibák.

5.4.1.1. Encoder

A hálózat a beágyazás részénél említett GNN-hez hasonló módon ország és szektor rétegekre rétegekbe van leképezve amik együtt képezik a neuron bemeneteket. A hálózat rejtett rétegeit továbbra is két gráf-konvolúciós réteg képezi, közöttük egy ReLU aktivációval, ami egy gyakran használt felépítés PyG-ben.

5.4.1.2. Decoder

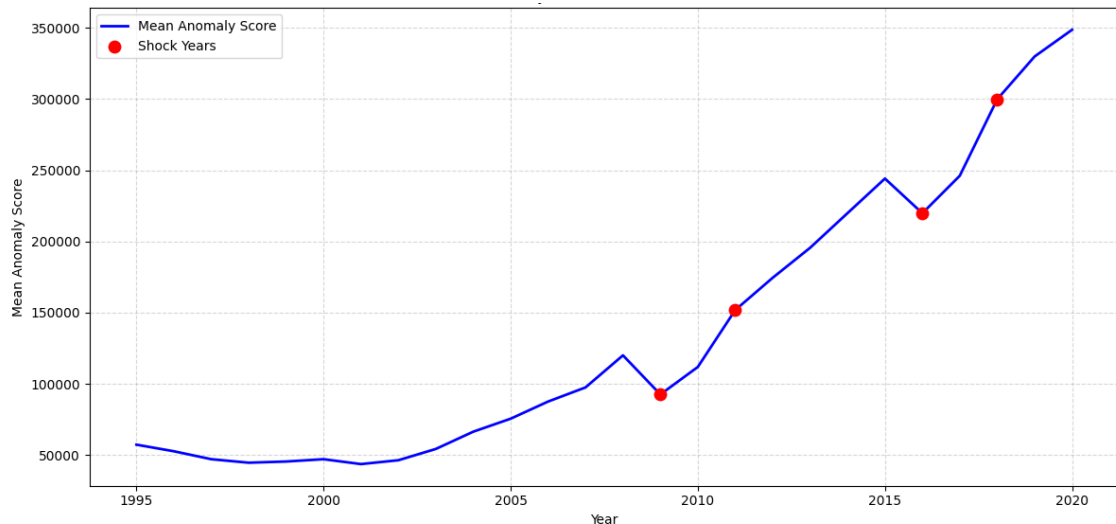
A hálózat “visszanyeréséhez” szükség van egy módszerre ami a hálózat éleit is rekonstruálja, ezt az implementált (és beszédes nevű) **reconstruct_edges** függvénnyel lehet, ami a csomópontok beágyazásának (pozíció-vektor) skaláris szorzata alapján becsüli meg az élek súlyát. Ez egy gyakran használt technika autoencoderekénél, a lényege az hogy ha két csúcs beágyazott koordinátái közel vannak, akkor köztük kapcsolat van.

5.4.1.3. Eredmények kiértékelése

Míg az első autoencoder megoldásom a rekonstrukciós hibát egy teljes gráfra számolta ki, itt a gráf-neurális megoldás esetén, a hálózat a gráf csomópontokra számolja ki a rekonstrukciós hibát, ami azt jelenti, hogy egy aggregációra van szükség, ami az évenkénti gráfokra meghatározza az adott év teljes rekonstrukciós hibáját.

5.4.1.4. Átlag szerinti aggregáció

Az átlag szerinti aggregáció esetén a rekonstrukciós hiba-vektor átlagát vettem ami egy átlagos értéket eredményezett minden csomópontra, ami így a komplett hálózat értéke lett. Ezeket az értékeket minden évre meghatároztam a gráfokból, ami diagramon megjelenítve a következő képet adta.



ábra 23: Rekonstrukciós hibák, átlag szerint aggregálva

Egy érdekes jelenség, hogy a globális kereskedelem folyamatos bővülése miatt megfigyelhető egy egyértelmű lineáris növekvő trend, ami a korábbi autoencoder megoldásnál az egyik mutatóból adódó zavaró tényezőként mutatkozott meg ezzel ev U alakú görbét létrehozva.

A hálózat által detektál sokkok itt inkább a mély, nagyobb volumenű recessziók. Ez az aggregáció képes megállapítani az egész hálózatot mélyen érintő változásokat. Azonban ebből fakadóan kevésbé képes detektálni a kisebb sokkokat a hálózatokban.

A modell sikeresen detektálta a 2008-2009-es világválságot és egy 2016-os visszaesést, amit okozhatott a Brexit, vagy a Trump első elnökségével induló első vámháború.

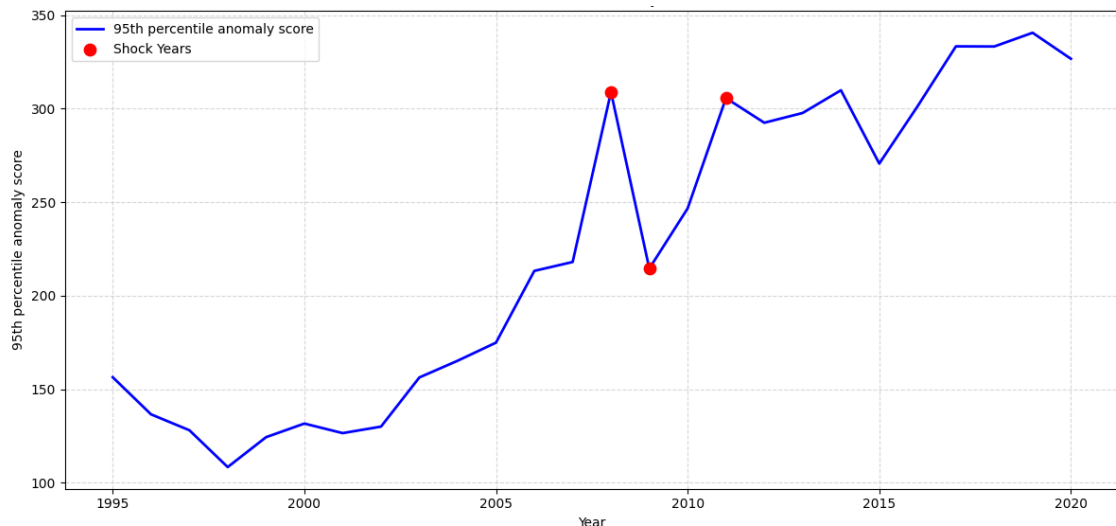
A grafikon két másik évek is sokként jelöl meg, azonban ez a trend miatt valószínűleg inkább a gazdasági fellendülésből adódó anomália. A 2011-es anomália köszönhető a világválságból való kilábalás utáni a nemzetközi kereskedelem visszaállása eredeti állapotára és a még ezen túli rohamos bővülésének. A 2018-as, pedig a 2016-tól tartó növekedési hullám csúcspontja.

Érdekes jelenség, hogy a gazdasági fellendülésből származó évek a visszaesést követő évek után 2 évvel következnek be.

5.4.1.5. 95. percentilis szerinti aggregáció

A 95. percentilis aggregáció ezzel ellentétben természetéből fakadóan érzékenyebb a szélsőséges értékekre, így jobban képes azonosítani a lokálisan jelentkező, de intenzív eltéréseket. Mivel nem az átlagos vagy medián viselkedést követi, hanem a legmagasabb 5% körüli anomáliaértékekre koncentrál, felfedi azokat az éveket, amikor bizonyos

ország-szektor párok szokatlanul viselkedtek, még akkor is ha a teljes rendszerben nem történt átfogó változás. Ez főleg akkor igazán hasznos, ha csak bizonyos szektorokat vagy országokat érintő változásokat vizsgálunk (például a szankciók, technológiai áttörés okozta változások vagy a 2025-ös vámháború érintette Európai autóiipar).



ábra 24: Rekonstrukciós hibák 95. percentilis mutatóval

Az ábrán jól látható, hogy a 2008-2009-es világválság egyes országokat, illetve szektorokat sokkal súlyosabban érintett mint másokat, míg ez a 2015-ös évet érintő válság esetén kevésbé volt jellemző. Ezek a kis, kevés szektort (vagy országot) érintő sokkok kereslet vagy kínálat eredetűek lehetnek, amit azonban sajnos az adathalmazon nem lehet vizsgálni.

5.4.1.6. A két megközelítés összevetése

A két módszer együtt alkalmazva átfogóbb képet ad arról, hogyan reagált a globális kereskedelmi hálózat az időben fellépő gazdasági és politikai folyamatok okozta sokkokra. Az egyik a „tájkép” változását figyeli, a másik pedig a „részleteket”.

6. A programkódok elkészítése

6.1. GPU használat

Az elemzések során a teljesítmény növelése érdekében GPU-t használtam, amit a NetworkX beépített CUDA támogatása tett lehetővé, a kódot `-use-cuda` argumentummal meghívva a NetworkX néhány függvényt az Nvidia videokártyámon futtatott. Mint

említettem ezt csak bizonyos függvények esetében tudtam alkalmazni, amelyek CUDA kompatibilisek voltak.

Továbbá mivel sajnos a dolgozat írásának idejében a CUDA támogatás windows rendszereken még nem volt natívan támogatva a python környezetet WSL-ben futtattam.

6.2. Függvények tesztelése

Az adatok átalakításának helyességét manuálisan ellenőriztem és az esetleges hibákat javítottam. A kódrészletek kiértékeléséhez automatikus tesztek nem készültek, ezeket manuálisan ellenőriztem. A kódot verziókezelve írtam és követtem git-ben, a repo elérhető GitHub-on. A kód dokumentációját LLM-ek (Large Language Model – Nagy nyelvi modell) segítségével végeztem, a könnyebb olvashatóság végett.

6.2.1. Reprodukálhatósági teszt (Reproducibility testing)

A tesztelés során a véletlen seed értéket változtatva is vizsgáltam a függvények kimenetét, a használt értékeket pedig a kódban feljegyeztem a reprodukálhatóság érdekében. A kód kész állapotban érhető el, azaz bármely részének futtatása az általam bemutatott ábrákat és értékeket mutatja, minden függvény reprodukálható a beállított seed értékeknek köszönhetően.

6.2.2. Ablációs tesztelés (Ablation testing)

Az „ablációs” vagy levételi, eltávolításos tesztelés lényege, hogy a függvény tesztelésekor a függvény bizonyos részit eltávolítva teszteljük a kapott eredményt. Ezt a megközelítést alkalmaztam az első, kezdetleges autoencoder megoldás implementálása során. A bemeneti réteget képező centrálítási mutatókat ugyanis a kimenet helyessége szerint választottam ki, ami végül a foksám centralitás a súlyozott foksám centralitás, lokális klaszterezési együttható, központiági centralitás, PageRank, eigenvektor centralitás és a k-mag szám lett. Sajnos azonban a tesztelés a hosszú lefutási idő miatt csak limitált kombinációval történt meg, azonban ez hozta a legjobb eredményeket.

Irodalomjegyzék

- [1] J. N. Pieterse, „Globalization and Culture: Global Mélange,” *Rowman: Littlefield*, 2015.
- [2] W. Y. Z. & L. J. Hamilton, „Inductive representation learning on large graphs,” *Advances in neural information processing systems*, 2017.
- [3] T. N. & W. M. Kipf, „Variational graph auto-encoders,” 2016.
- [4] K. B. A. J. & D. M. Panford-Quainoo, „Bilateral trade modelling with graph neural networks,” *In ICLR workshop on practical ML for developing countries.*, 2020.
- [5] Z. T. D. K. a. B. F. Kosztyán, „Trade network dynamics in a globalized environment and on the edge of crises,” *Journal of Cleaner Production*, 2024.
- [6] G. Szini, „Correcting for Sample Selection Bias in Dyadic Regressions,” *Diss. University of Amsterdam*, 2020.
- [7] P. Vakhil, „The evolution of downstream global value chains in Eastern Europe,” *Society and Economy*, 2023.
- [8] H. I. Yasuyuki Todo, „Propagation of economic shocks through supply chains,” *VoxEU*, 2019.
- [9] V. S. Isabelle Rosenberg, „On Shock Propagation,” *MSc Thesis*, 2018.

Mellékletek

Mappaszerkezet:

```
+ misc
|   +---centrality_measures
|   |   +---multilayer
|   |   |   __init__.py
|   |   |   calculate_indicators.py
|   |   |   operations.py
|   |   |   visualizations.py
|   |   |
|   |   +---singelayer
|   |   |   __init__.py
|   |   |   calculate_indicators.py
|   |   |   operations.py
|   |   |   visualizations.py
|   |   __init__.py
|   |
|   +---community_detection
|   |   +---multilayer
|   |   |   __init__.py
|   |   |   calculate_indicators.py
|   |   |   operations.py
|   |   |   visualizations.py
|   |   |
|   |   +---singelayer
|   |   |   __init__.py
|   |   |   calculate_indicators.py
|   |   |   operations.py
|   |   __init__.py
|   |
|   __init__.py
|   database_operations.py
|   localmemory_operations.py
|
main.py
```

Ábrajegyzék

ábra 1: Intralayer vizualizáció, az EU országairól három szektor kereskedelmével	16
ábra 2: Kontinensek szerinti fokszámközpontiság érték 1995-től (kék) 2020-ig (piros) 17	
ábra 3: Országokénti Fokszám központiság 1995-től 2020-ig	18
ábra 4: Közöttiségi központisági értékek 1995-től 2020-ig gyarmatosító országok esetében.....	19
ábra 5: Harmonikus központisági értékek (térképen) 1995-ban és 2020-ban	20
ábra 6: Top 50 Pagerank értékű szektor 1995-től és 2020-ig országok által csoportosítva	22
ábra 7: Denzitás értékek 1995-től 2020-ig, a teljes hálózatra	24
ábra 8: Top 10 ország rétegbeli denzitása 1995-től 2020-ig.....	26
ábra 9: Top 10 szektor rétegbeli denzitása 1995-től 2020-ig.....	27
ábra 10: Rétegek közötti denziás országoként 1995-től 2020-ig	28
ábra 11: Rétegek közötti denzitás szektoronként 1995-től 2020-ig.....	29
ábra 12: Homofília változása 1995-től 2020-ig	29
ábra 13: Louvain algoritmus által megállapított modulok 1995-ben.....	31
ábra 14: Louvain algoritmus által megállapított modulok 2020-ban.....	32
ábra 15: Mag-periféria által megállapított mag-határok 1998-ban	33
ábra 16: Jaccard távolság az évekre számított Louvain modularitásra	35
ábra 17: Sokkelemzés a DCD-nél azonosított értékekre	36
ábra 18: Autoencoder által megállapított sokkok 1995-től 2020-ig	38
ábra 19: Lokális maximák a Sokkelemzés során megállapított gráfon	39
ábra 20: Beágyazott eredmények V4-es országokra 2002-től 2014-ig (3 évente).....	41
ábra 21: Beágyazott csomópontok, az elektronikai ipar fontosabb szereplői.....	44
ábra 22: Klaszterek megállapítása K-means segítségével 2005-től 2009-ig	46
ábra 23: Rekonstrukciós hibák, átlag szerint aggregálva.....	49
ábra 24: Rekonstrukciós hibák 95. percentilis mutatóval	50

Táblázatjegyzék

<i>1. Klaszterek idősoros táblázata 2005-től 2010-ig</i>	<i>47</i>
--	-----------