Assignment 2

Militarized Interstate Disputes in a Postmodern World

Xuesi Hou & Thomas Haase

February 21, 2025

Table of contents

---

1 Task 1

Formulate theoretical arguments and expectations about interstate relations. Think in network terms. Select the appropriate part of the data for analysis.

1.1 Introduction to the Data

The selected dataset is the Military Interstate Disputes V5.0 dataset of the *correlates of war* project. It contains information about "all instances of when one state threatened, displayed, or used force against another."(Palmer et al. 2022). Each incident contains a description and is linked to all countries including who initiated the conflict, the start and end time, casualties and most importantly the countries connected to the incidents.

This bipartite graph of countries connected to incidents is projected to a network where only countries are connected to each other based on how many incidents they shared/were part of. We filter the data only for conflicts happening in the post modern world after 1949.

1.2 Expectations

It is not far fetched to theorize that geographical location is playing a major role in sharing military incidents, because those borders are most probable not only the result of such disputes but being a neighboring country also can pull a country inside of a conflict that it was originally no part of.

Since the data was filtered to include only the years 1949 - 2014, the analysis is expected to reflect the most important geopolitical conflicts of this era. The major terms under which many incidents happened were the cold war era (1949-1991) and the post-cold war era. During the cold war the US and Russia fought a series of proxy wars like:

- Korean War (1950-1953)
- Vietnam War (1955-1975)
- Cuban Missile Crisis (1962)
- Arab-Israeli Wars
  - Six-Day War (1967), Yom Kippur War (1973)
- Soviet-Afghan War (1979-1989)
- Iran-Iraq War (1980-1988)

The post-cold war era (1991-2014) is characterized by mostly wars in africa, the middle east and eastern europe:

- Gulf War (1990-1991)
- Yugoslav Wars (1991-2001)
- First (1994-1996) and Second (1999-2009) Chechen War
- Congo Wars (1996-2003)
- Afghanistan War (2001-2014)
- Iraq War (2003-2011)
- Arab Spring conflicts (2010-2014)
- Russo-Georgian War (2008)

- Syrian Civil War (2011-ongoing)
- Crimean Crisis (2014-ongoing)

Because of their hegemonial position in the world order the US, Russia and Europe probably have a very high degree (were part of many incidents). Many of the listed wars are local with the involvement of a few larger nations as perpetuators or mediators in war/military disputes. We expect the data to represent this relationship.

2 Task 2

> On the selected data, make a table that contains descriptive statistics for the network and for the distribution of key node-level variables. Please mention the historical period covered, the nature of tie(s), the number of nodes and ties, average degree, density, the standard deviation of degrees (or outdegrees and of indegrees if network is directed), and average path length. Identify the number of isolates if there are any. Please check if the network is a single component. Which node has the highest degree, and which has the highest betweenness centrality? What do these measures represent in terms of the specific type of relations and data you have selected?

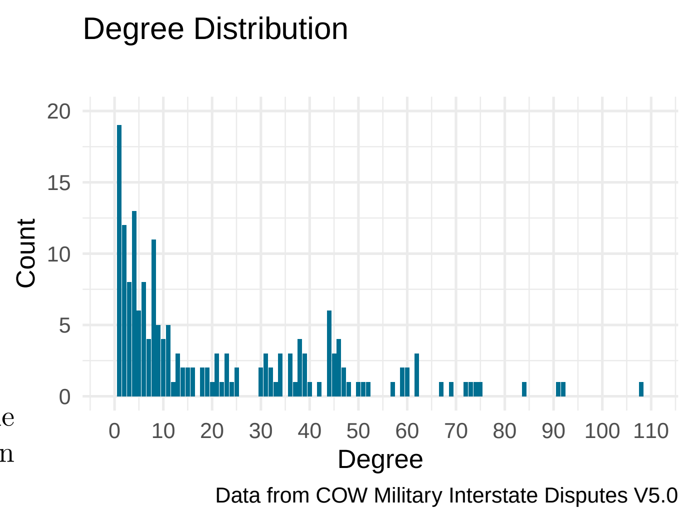
The analysis transforms a two-mode network (bipartite graph) into a one-mode network through projection. The original network connects two distinct types of nodes: countries and incidents. In the projected network, only countries remain as nodes, and they become connected when they were involved in the same incidents. The strength of connection between any two countries is determined by the number of shared incidents between them. The analysis focuses exclusively on the contemporary post-World War II era, specifically from 1949 onwards until 2014, the last year recorded in the dataset.

Descriptive Summarystatistics

Military Interstate Disputes[1]

Parameter	Value (AVG ± SD)
Node Count	181
Tie Count	1964
Degree	21.7 ± 22.7
Average Path Length	6.9
Component Count	1
Component Size	181

[1]Countries are connected by being part of the same dispute and weighted with the count of occuring in shared disputes.



Top Countries by Centrality Measures[1]

Top Degree	Degree Score[2]	Top Betweenness	Betweenness Score[2]	Top Eigenvector	Eigenvector Score[2]
USA	108	USA	0.16	AZE	1.00
UKG	92	UKG	0.10	ARM	1.00
FRN	91	FRN	0.10	RUS	0.06
RUS	84	IRN	0.08	TUR	0.06
NTH	75	LAT	0.07	USA	0.05

[1]Countries are connected by being part of the same dispute and weighted with the count of occuring in shared disputes.
[2]Scores are calculated including weights

The table with descriptive summarystatistics describes the network having 181 nodes, including most existing countries in the postmodern world (since 1949). The network has 1964 ties, so at a network level without considering nodedegree we would expect each node to have $\frac{1964}{181} \sim 10$ connections. So on average a node is involved in military disputes with ~ 10 other countries.

The average path length being 21.7 shows that on average a country is connected to another country via ~20 disputes on average. What is striking is the very large standard deviation of degrees. The Standard deviation of the degree measure is larger than the average itself, indicating the existance of massive superhubs. This observation is backed up by the plog of degree distributions showing a very long tail. The plot of degree distributions reveals many countries having degreecount of less than 20, which means most countries are part less than 20 incidents. There is a smaller spike at degreecount 50, which means that there seems to be a set of countries involved in many more incidents. These could also be the result of some incidents involving the same countries many times. One country is a big outlier with over 100 degrees.

The average path length of the network is 6.9 and therefore very short. This implies the existance of shortcuts, similar to the small world network (Watts 1999). This could mean that there are military incidents connecting countries from different hubs.

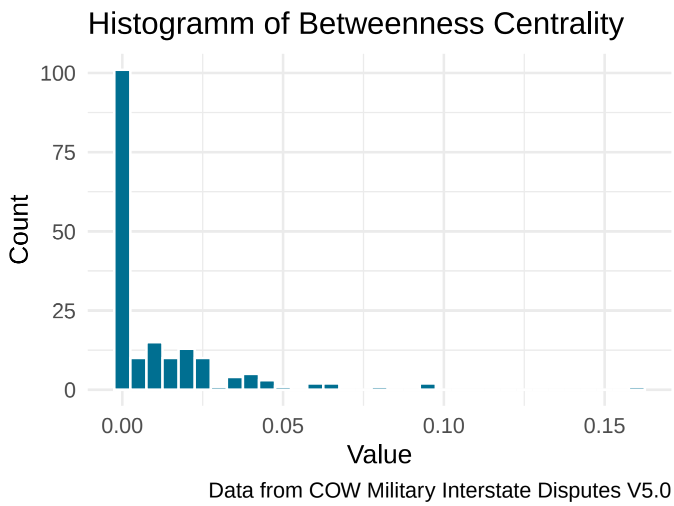
The network consists of only one comoponent due to only including countries that are actually linked by military interstate desputes.

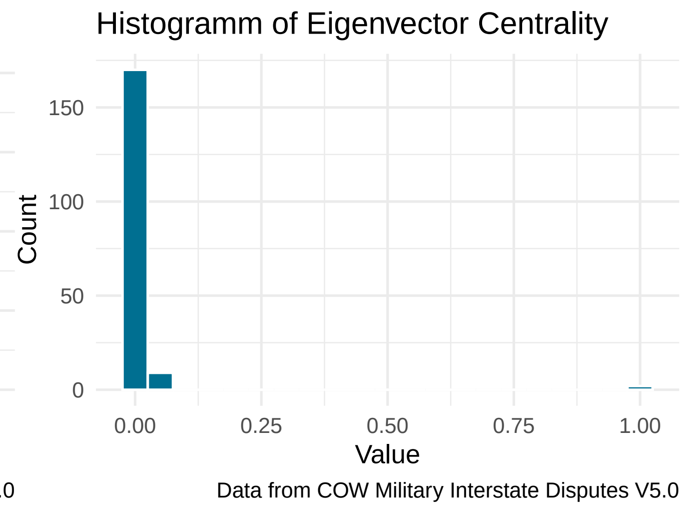
The table "Top Countries by Centrality Measures" presents centrality measures for countries involved in military interstate disputes. The United States leads with the highest degree score of 108, followed by United Kingdom (92), France (91), Russia (84), and the Netherlands (77). This indicates that the US was involved in the most shared military incidents during this period, reflecting its global military presence and engagement in numerous conflicts.

The betweenness centrality scores show the US again leading (0.16), followed by United Kingdom (0.10), France (0.10), Iran (0.08), and Latvia (0.07). These scores suggest these countries often served as critical intermediaries in the conflict network, potentially playing important roles in regional disputes or international interventions.

Azerbaijan leads with a score of 1.00, followed by Armenia (1.00), Russia (1.00), Turkey (0.06), and the US (0.05). The high eigenvector scores for Azerbaijan, Armenia, and Russia suggest these countries were connected to other highly central actors in the network, particularly in regional conflicts.

This data aligns with historical events during this period. The US's high degree and betweenness scores reflect its involvement in major conflicts like the Korean War, Vietnam War, and Gulf Wars. The prominent positions of United Kingdom and Russia align with Cold War dynamics and post-Soviet conflicts. The high eigenvector scores for Caucasus region countries (Azerbaijan, Armenia) and Turkey indicate the significance of regional conflicts in that area. The presence of both Western powers (US, France) and Eastern powers (Russia) at the top of these rankings reflects the bipolar nature of the Cold War period and its aftermath, demonstrating how these powers were central to global military disputes during this era.





The Betweenness Centrality Histogram shows a highly skewed distribution with the vast majority of countries having very low betweenness centrality. There's a steep drop-off after the first bar, indicating that very few countries serve as important "bridges" in the network. Only a handful of countries (as we saw in the previous table - like USA, United Kingdom, France) have higher betweenness scores around 0.10-0.15. This suggests that most countries are involved in regional or localized conflicts, while a small number of powerful nations act as intermediaries across different conflict zones

The Eigenvector Centrality Histogram shows an even more extreme concentration, with almost all countries having very low eigenvector centrality scores of 0. There's a small group of countries with very high scores (around 1.0), which we know from the table includes Azerbaijan, Armenia, and Russia. The stark binary nature of this distribution (mostly either very low or very high) suggests a clear hierarchy in the conflict network, there seem to be some disputes connecting very "popular" countries (countries involved in many disputes) to Azerbaijan, Armenia and Russia.

Together, these distributions reveal a highly centralized network structure where a small number of countries play disproportionately important roles in global military disputes, while the majority of nations have more limited involvement. This aligns with what we might expect given the global power dynamics during the Cold War and post-Cold War periods, where superpowers and regional powers dominated international military engagements.

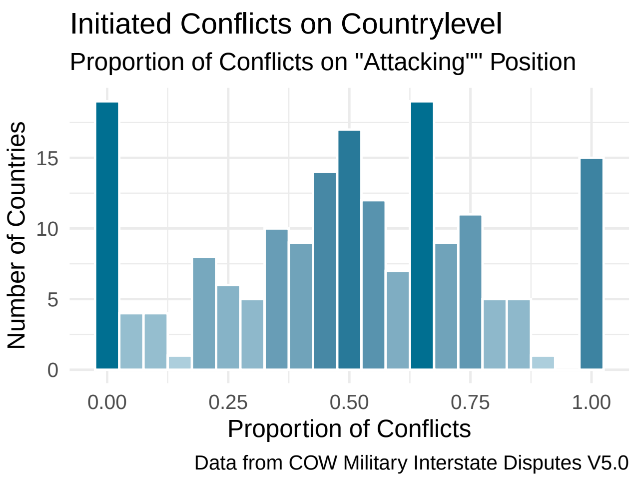
Top Countries by Fatalities Count

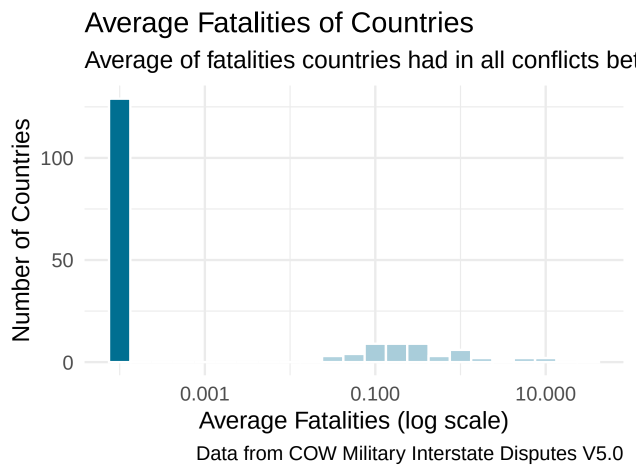
Country	Fatality Count
Iraq	7025
Pakistan	950
Azerbaijan	684
India	562
Armenia	434

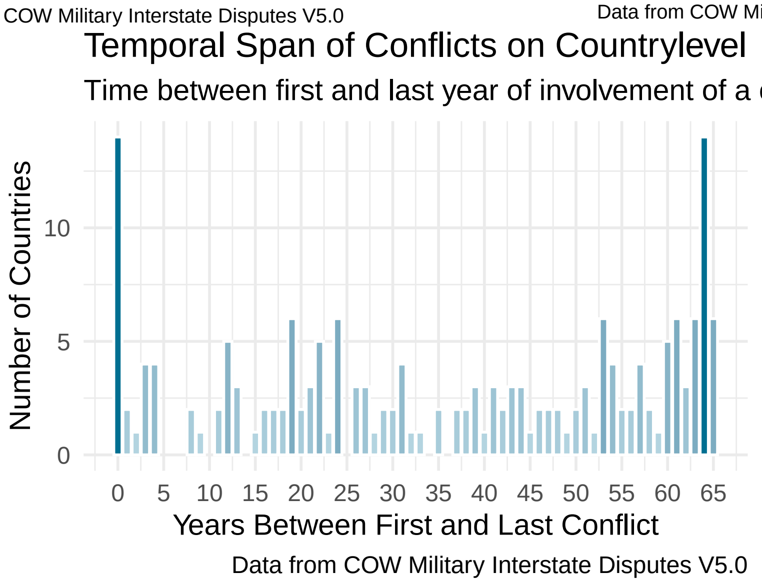
3 Rewrite Interpretation

The table about the Countries with most fatalities in the dataset is best interpreted using the description of the disputes of the certain countries.

Iraq's count of 7,025 fatalities stands markedly higher than other countries, indicating its involvement in major conflicts including the Iran-Iraq War and Iraq War. Pakistan and India's figures (950 and 562 fatalities respectively) stem largely from their Kashmir territorial dispute, while Azerbaijan and Armenia's numbers (684 and 434) reflect intense regional conflicts in the Caucasus following the Soviet Union's dissolution.







The initiated conflicts histogram reveals that countries typically initiated between 25-75% of their military engagements, with a notable peak around 50%. This distribution implies that most nations alternated between initiating and responding to military disputes, rather than consistently assuming either role.

The fatalities histogram demonstrates a right-skewed distribution, with most countries experiencing relatively low average casualties while a small number of conflicts resulted in significantly higher fatality rates. This pattern aligns with the presence of both limited military engagements and more intensive warfare during the studied period.

The temporal span distribution indicates considerable variation in conflict duration, with most disputes resolving within 20 years while some extended conflicts persisted for up to 40 years. This bimodal distribution suggests the coexistence of both acute military engagements and protracted disputes during this period.

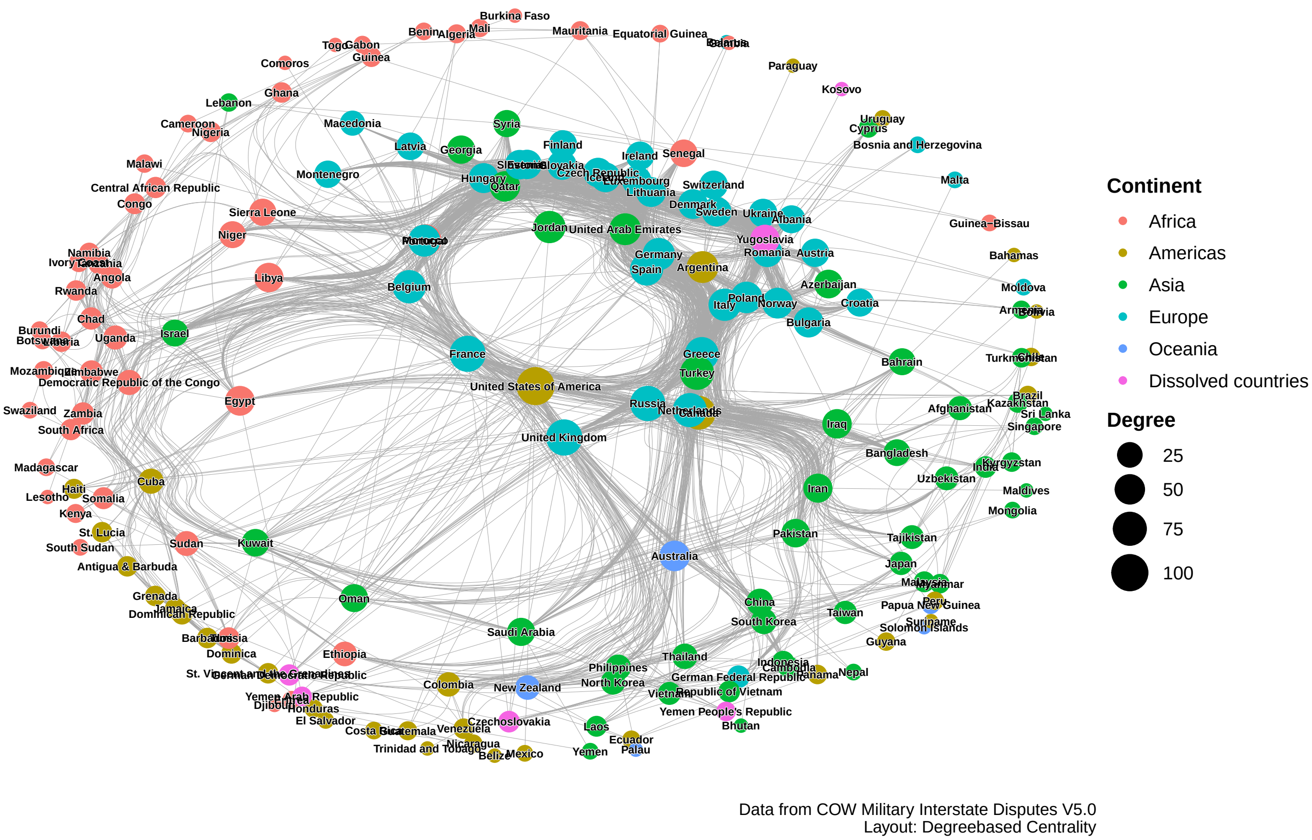
The analysis reveals three main patterns in military conflicts: First, most nations showed balanced engagement, initiating about half of their disputes. Second, while most conflicts had relatively low casualties, a few resulted in significantly higher fatalities. Finally, conflict durations varied widely, with most lasting under 20 years but some extending up to 40 years.

4 Task 3

Visualize the network. Vary degree color and size according to key node attributes. Try to pick a node attribute that relates to an important characteristic (e.g., power, GDP per capita).

Countries connected by shared Military Interstate Disputes

By Continents



The network shows European nations are densely connected, particularly in Eastern Europe, indicating frequent military interactions within the region. This pattern likely reflects both Cold War tensions and post-Soviet conflicts. Asian countries show strong regional clustering, with notable connections centered around key regional powers and territorial disputes. African nations appear more dispersed throughout the network, suggesting more localized conflicts rather than continent-wide military engagements. The Americas display relatively fewer interconnections, with most countries maintaining lower degree centrality except for the United States. The United States dominate the network with their high degree being part of most conflicts. The presence of dissolved countries adds an important historical dimension, particularly in Eastern Europe, highlighting how geopolitical restructuring influenced military dispute patterns. Especially Yugoslavia being in the central cluster of european states highlights the power of the now not anymore existing state.

Especially interesting is the representation of large international conflicts with a lot of parties by the bundled edges. Especially visible is this phenomenon with Iraq, Iran, Pakistan and Isreal.

5 Task 4

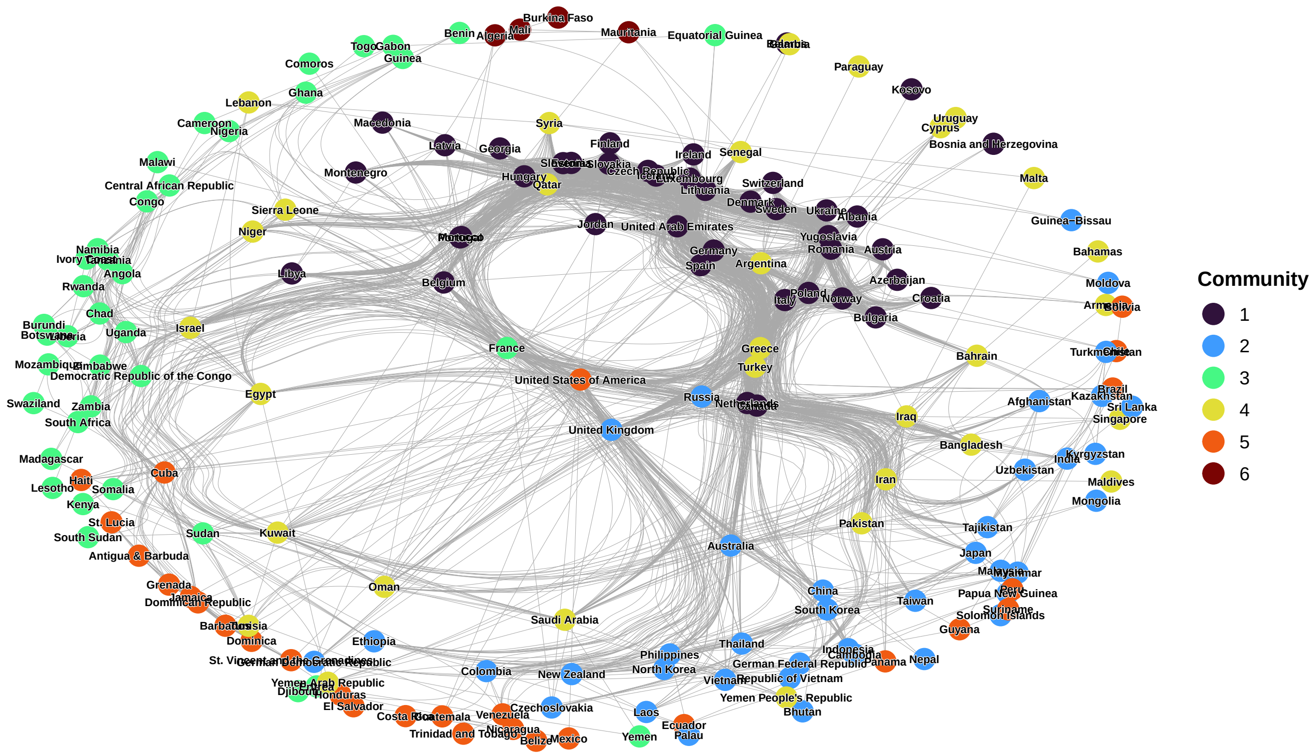
> Implement and compare the results (community compositions, their size) of TWO different community detection algorithms to find communities (e.g., Girvan-Newman, Louvain, walktrap, cohesive blocks, Leiden). Please take into account that most of these methods can use undirected graphs only, others (such as Girvan-Newman) might work with weighted graphs. Before carrying out this task, simplify the data if necessary. Create a plot where nodes are either colored based on detected community membership, or you can circulate the groups with some color in the background.

The Louvain algorithm visualization identifies five communities. They represent mainly regional grouping from the continents, but some countries in the center regions now are part of the outer regional communities. For example the France is now part of a cluster with some African countries and countries from the middle east. Some Western Asian and Middle eastern countries are now part of the European cluster. The Louvainclustering therefore revealed partly geographical distinctions in Asia, aswell as hegemonial and power-related structures (regarding the middle east), aswell as highlighting some major conflict involvements (Frances position).

The Spectrality based Leading-Eigenvector Clustering revealed 8 communities. The communities are very similar to the results of the Louvain clustering. One difference between the clustering algorithms would be that the African countries are now differentiated between those connected to military disputes with France, while central Asia is also their own community (Tajikistan, Mongolia, Uzbekistan,...).

Community Detection: Louvain

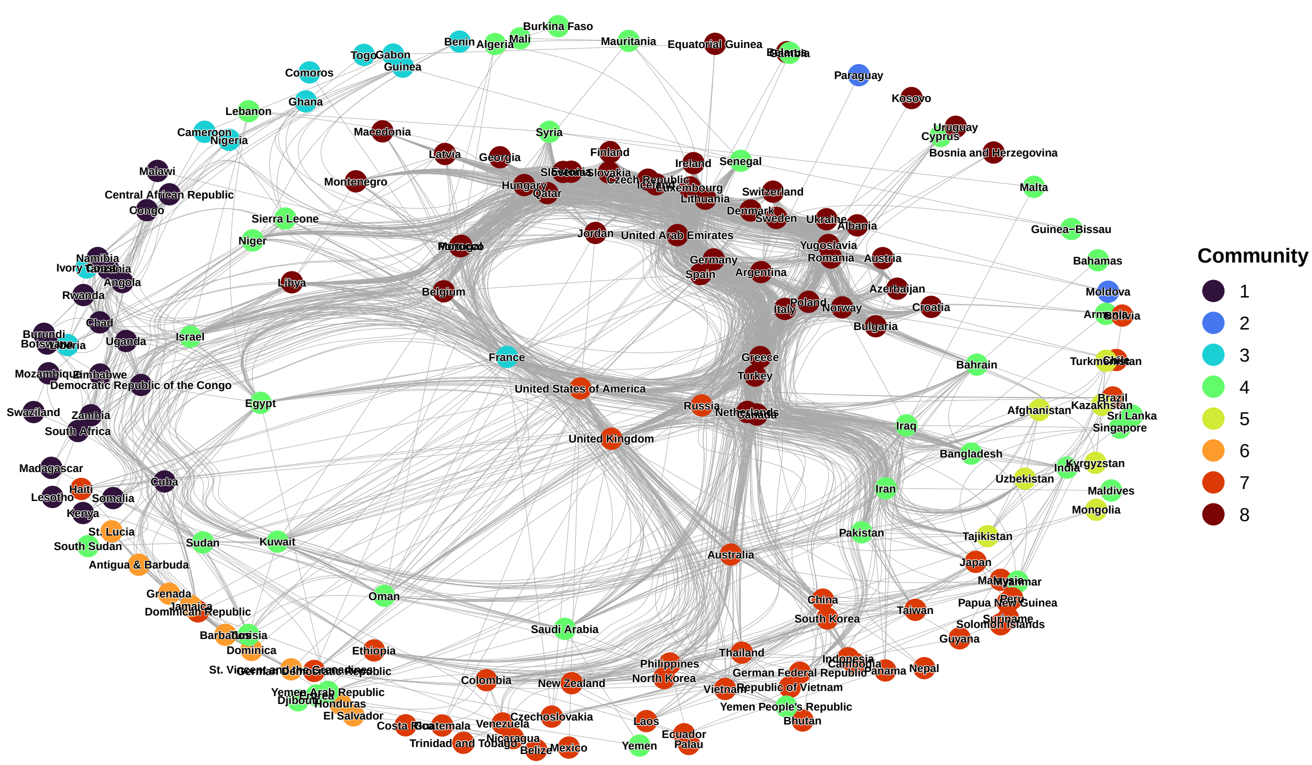
Countries connected by shared Military Interstate Disputes



Data from COW Military Interstate Disputes V5.0
Algorithm from Blondel et al. (2008)
Layout: Degreebased Centrality

Community Detection: Spectrality based Leading−Eigenvector

Countries connected by shared Military Interstate Disputes



Data from COW Military Interstate Disputes V5.0
Algorithm from Newman (2006)
Layout: Degreebased Centrality

6 Task 5

> Make meaningful comparisons between the results of community detection algorithms. Summarize the differences and similarities and try to explain these. Please refer to the relevant literature describing the logic of these algorithms if necessary.

As mentioned in Section 5 both algorithms indicate geographical aswell as powerrelated structures of the world order. Some conflicts and colonial relations are large enough to push through this geographical pattern to form and influence the grouping of communities.

6.1 The Louvain Algorithm

The Louvain Algorithm is a greedy technique for clustering a network into non-overlapping communities (Blondel et al. 2008). At first, each node is their own community. Then the nodes are clustered locally via modularityoptimization. After reaching the first local maxima the algorithm proceeds with phase 2 which aggregates the communities into blocks including elements of a weighted graph, with edge count between them. Then clustering is repeated.

6.1.1 Intuition

In case of the used data countries are connected by shared military disputes. Two countries are part of the same community if they shared considerably more disputes with each other than with other countries (=modularity-optimization). These local clusters of common friends and enemies are used in the second step to optimize the clusters even further. By this we reach clusters of countries who frequently interact in the context of military incidents.

6.2 Spectrality based Leading-Eigenvector Algorithm

Newman (2006) developes a clustering algorithm which is based on the usage of a modularity matrix allowing a leading-eigenvector based optimization elevating spectral techniques (Newman 2006).

"The complete iterative algorithm comprises of two steps. In the first step, the network is expressed in terms of its leading eigenvalue and eigenvector and recursively partition into two communities. Partitioning occurs if the maximum positive eigenvalue is greater than the tolerance threshold for the current partition, and if it results in a positive contribution to the modularity. Given an initial separation using the leading eigen step, the function then continues to maximise for the change in Modularity using a fine-tuning step [...]. The first stage is to find the node which, when moved from one community to another, gives the maximum change in modularity. This node's community is then fixed and [the process is repeated] until all nodes have been moved. The whole process is repeated from this new state until the change in the Modularity, between the new and old state, is less than the predefined tolerance."(Mclean and Sorokin 2022)

6.2.1 Intuition

The modularity matrix contains elements $B_{vw} = A_{vw} - \frac{k_v k_w}{2m}$ representing the difference between the expected and observed tie for each observed edge (Newman 2006). In terms of our network those values are the difference between the countries we expect to share military incidents and the countries that actually share military incidents. This depicts the notion of a country to connect to another country via shared military incidents.

The eigenvector of this matrix describes the orientation of distortion that countries notions to connect to each other through shared military disputes in the space of all countries. The eigenvalue describes the scale into the direction of said distortion. The exact same procedure is used for Principal Component Analysis and this similarity is even mentioned by Newman (2006).

A community can then be expressed as the orthogonal vector of the vector of edgecombinations which is similar to the result of PCAs. This orthogonal vector then cuts the space, which describes the deviation from expected and observed edges, in 2. This cut is chosen so it optimizes modularity. In terms of our dataset: The communityvector should in theory describe the incidents which were part of the bipartite graph before we projected the bipartite graph to a country network. Following this, the orthogonal vector of the incident-vector cuts modularitymatrix $B_v w$ in 2 large clusters optimizing modularity so that the result are two super-groups of incidents splitting the country-projection with the optimal modularity. This is splitting/"PCA-on-graphs" is repeated until a certain tolerance threshhold is reached.

6.3 Comparison

In conclusion both algorithms in theory should be able to retrieve the most important information of incidents which got lost during the projection of the bipartite graph. They also implicitly depict the underlying structures generating the edges of being part of the same international military disputes. These underlying structures seem to be a structure of international hegemony still reproducing structures generated by the colonial era aswell as geographical location and sharing borders.

Bibliography

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." *J. Stat. Mech.* 2008 (10): P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Mclean, Colin, and Anatoly Sorokin. 2022. "rSpectral: Spectral Modularity Clustering." https://doi.org/10.32614/CRAN.package.rSpectral.

Newman, M. E. J. 2006. "Finding Community Structure in Networks Using the Eigenvectors of Matrices." *Phys. Rev. E* 74 (3): 036104. https://doi.org/10.1103/PhysRevE.74.036104.

Palmer, Glenn, Roseanne W McManus, Vito D'Orazio, Michael R Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, and Michael J Soules. 2022. "The MID5 Dataset, 2011–2014: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 39 (4): 470–82. https://doi.org/10.1177/0738894221995743.

Watts, Duncan J. 1999. "Networks, Dynamics, and the Small-World Phenomenon." *American Journal of Sociology* 105 (2): 493–527. https://doi.org/10.1086/210318.