THE UNIVERSITY OF CHICAGO

Computational Analysis of the United Nations General Assembly Resolutions: Prevailing Issue Areas and International Alliance Formation

By

Egemen Pamukcu

June 2022

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: Brooke Luetgert
Preceptor: Shilin Jia

I. Introduction

A. Background

The United Nations General Assembly (UNGA) was established with Chapter IV of the United Nations (UN) Charter in 1945 based on the principles of universal membership and sovereign equality. Since its founding, the UNGA has been the primary platform in which all the UN representatives from all member countries convene and discuss the world's pressing issues concerning a range of policy areas such as world peace, environment, and human rights, as well as some administrative and financial matters essential to the functioning of the UN. As of March 20, 2022, the UNGA has 193 voting members and has passed 19,433 resolutions, 6,895 of which were voted upon. A simple majority is required for a resolution to pass, however, UNGA resolutions, unlike those of the UN Security Council (UNSC), except for some extraordinary situations and budgetary issues, are not binding on its members—Articles 10 and 14 of the UN Charter refers to the UNGA resolutions as "recommendations." Although there are many critiques of the UNGA centered around its fairness and effectiveness (Easterly, 2009) (Holmes, 1993) it still remains as the main deliberative organ of the most representative international organization to date.

B. Motivation

There have been numerous attempts by international relations scholars to come up with a model of the global order and primary sources of international conflict both before and after the Cold War period (Sakwa, 2019). One overarching example is the Clash of Civilizations hypothesis proposed by Samuel Huntington which predicted that the post-Cold War interstate conflicts will be driven by cultural differences between the nine "major civilizations" such as the Western, Sinic, and Islamic civilizations (Huntington, 2000). Another widely accepted macro-alliance structure, especially within international development and international political economy (IPE) circles, is the even broader distinction between the countries of Global North and the Global South (Lees, 2020). This grouping is often brought up in discussions of international trade and sustainable international development. The argument is often that developing countries of the Global South, in order to improve living standards, need to prioritize national industrial development over trade liberalization or environmental concerns such as high CO2 emissions, and that the Global North is imposing these standards on the South after achieving easier-to-maintain scientific know-how and industrial infrastructure in the past when these "international standards" were not yet in place (Kacowicz, 2007).

These higher-level groupings are perhaps useful in providing a framework to think about and discuss the most pressing issues within the field of international relations, however, most of them suffer from three major shortcomings. Firstly, these frameworks, by providing a useful and easy to comprehend country blocs to think about international political issues, sacrifice precision. In other words, often African and South Asian countries are grouped in one "bloc", or European countries are conceived to be acting in perfect synchronization. When it comes to real policy discussions, we suspect that these blocs tend to be overgeneralized and there can be systematic outliers within these groups, i.e. one country can be systematically defying their assigned bloc which may lead to an incomplete or a noisy understanding of the issue at hand (Kloß, 2017). Secondly, these traditional groupings are often all-encompassing in terms of issue areas and do not acknowledge that two given states can be adversarial when it comes to, for instance, extraction of marine resources but highly cooperative in the field of international human rights. A universal, all-encompassing structure for the international order often falls short in capturing this intertwined and multidimensional nature of international politics. Finally, these groupings are often slow in adapting to changes in interstate dynamics across time. The nature of interstate alliances can change with, among many things, developments in national politics (e.g. election of a populist government), major global events (e.g. economic or human rights crises, a global pandemic), or black swan events such as unforeseeable escalation regarding a disputed territory. Frameworks such as the First, Second, and Third World still used after the Cold War, do not hold well against the test of time and lose explanatory power every passing year (The World Bank, 2010).

There have been several efforts to use empirical data to conceptualize international alliance dynamics. Macon et al. (2012) studied the UNGA community structures of networks representing the votes on UNGA sessions by considering considering voting similarities as weighted, unipartite networks. Traag and Bruggeman (2009) applied an extended version of the Potts model for community detection on signed networks to the Correlates of War data set over the 1993-2001 preiod and found six power blocs corresponding roughly to the Huntington's civilizations. Voeten (2000), applying the a multidimensional scaling technique (NOMINATE) on the UN voting data, investigated the determinants as well as the stability of voting alignments during pre- and post-Cold War periods, finding fairly stable blocs spread across a single dimension mostly characterized by wealth and democracy, which closely resembles the Cold War East-West dimension with few exceptions. The network approach to understanding community structures has been applied to many other areas within the social sciences as well. Porter et al. (2005) employed methods of network analysis and hierarchical clustering to analyze the committees in the U.S. House of Representatives. Whereas, Waugh et al. (2011) examined the party polarization in the U.S. Congress through network modularity and investigated “divisiveness” and “solidarity” measures.

This study will aim to provide an empirical framework to offer an alternative approach to analyzing and understanding international alliance structures that is sensitive to granularity and multidimensionality both along the axis of time as well as issue areas. Through the voting data on the UNGA resolutions I extract the higher-level country clusters using unsupervised machine learning techniques to reduce dimensionality and cluster the groups. Later, the Natural Language Processing (NLP) technique of topic modeling is applied on the textual content of UNGA resolution documents to extract topics that correspond to issue areas discussed on the floor of the General Assembly. And, finally, I use graph data structures and their network visualizations to get an empirical understanding of international alliance structures across these extracted topics. This combination of methodologies, by taking the textual content of the UNGA resolutions into account, allows me to compare and contrast alliance structures across time and issue areas, visualize these complex network structures to facilitate their qualitative analysis, and, ultimately, answer the questions of to what extent these alliance networks reflect (i) the higher-level clusters extracted from the voting data that does not take document topics, and as a result issue areas, into consideration or (ii) the classical and broadly defined frameworks such as the North-South, West-East, or cultural civilizations distinctions.

II. Data

A. Data Collection and Generation

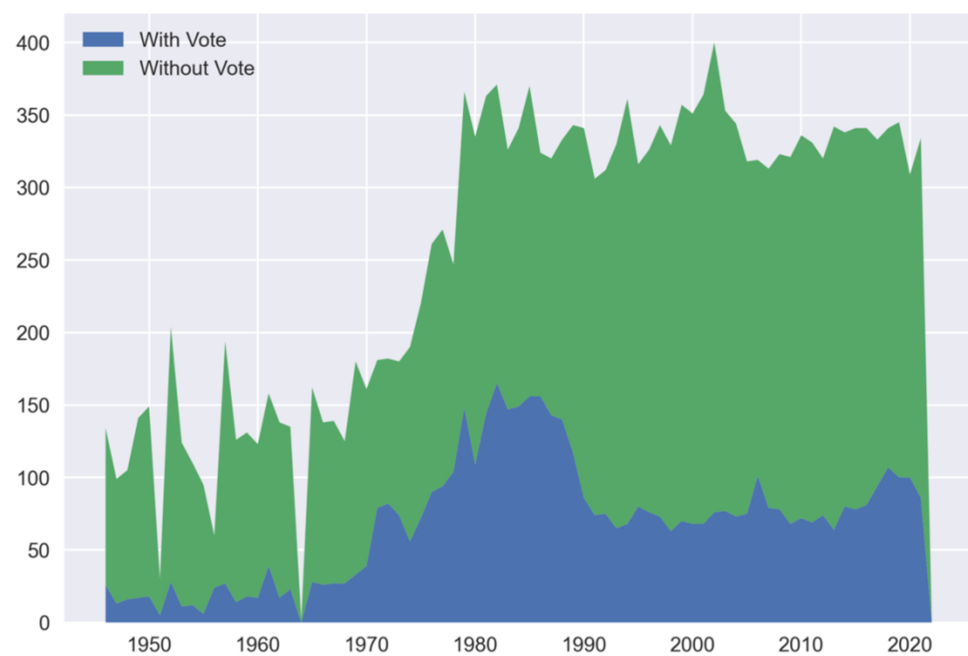
One of the main contributions of this study is making a corpus of UNGA resolutions along with the relevant metadata and roll call records publicly available in a machine-readable format. For this study, the HTML content of the United Nations Digital Library System was scraped to acquire data and metadata on 19,102 UNGA resolutions. The resolution ID, title, date, voting tally (where applicable), and the URLs directing to XML and PDFs containing the individual voting records and resolution text were captured in the first pass through the UN Digital Library website. In the second pass, the voting dataset was created with rows as resolutions, columns as countries, and decisions in individual cells. In total, there are five possible values in voting data. A country can choose to vote "Yes", "No", or "Abstain" for the motion, can choose not to attend the voting session, may not exist at the date of voting, or may not be a UN member at the date of voting. The last two categories are accepted as NA (missing) values. In the third and final pass, the text content of the PDFs was collected and converted into a machine-readable format.

Especially for older resolutions, most of the PDFs contained scanned images of physical documents instead of actual text. Google's open-source Tesseract Optical Character Recognition (OCR) Engine was used to convert the images into the string format (Smith, 2007). Some of the older English documents also contained French translations side-by-side in the same document. To handle those cases a language detection API was used to filter out the sentences in other languages. Also, possibly due to a bug in the UN system, for 25 resolutions the only available language was German and all URLs for different languages directed to the German version of the resolution—these were left out of the corpus. A small portion of the URLs in the UN system either directed to corrupted PDF files or were missing altogether. As a result, they were not included in the analysis. I must note that a manual examination of a sample of the missing resolutions revealed no systemic patterns that could have a discernible impact on the findings.

B. Data Facts

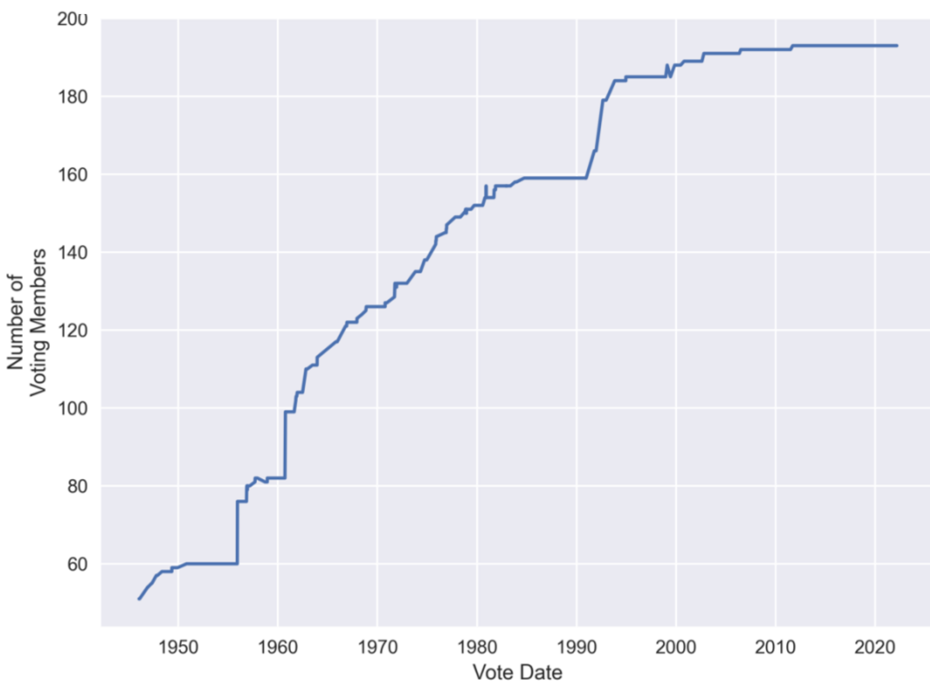
After collection, cleaning, and pre-processing of the data, the final dataset contains metadata on a total of 19,428 resolutions. 5,235 of these resolutions were accepted as a result of voting among the member countries.

Figure 1: Number of resolutions adopted by the UNGA between 1946 and 2022



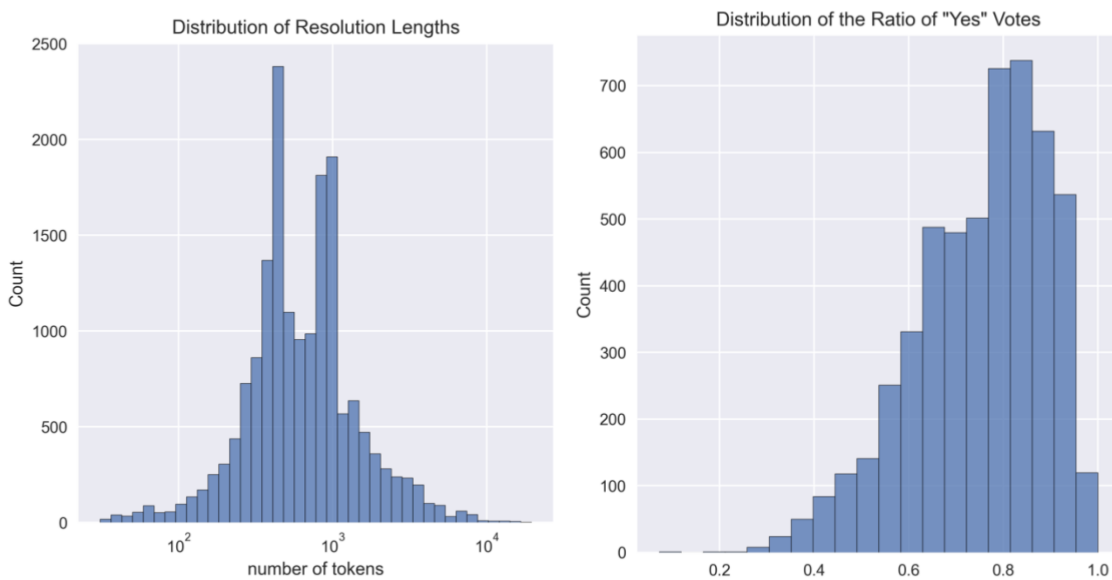
After accounting for corrupted and missing files, the corpus contains a total of 17,199 documents, resulting in a completeness rate of 88.5%. The first resolution that is included in the corpus was adopted on January 24th, 1946, and discusses the terms for the appointment of the Secretary-General of the United Nations. The latest resolution in the dataset was adopted on March 2nd, 2022, and discusses the recent Russian aggression against Ukraine. Since its founding, a total of 202 unique countries have participated in UNGA voting until March 2022, with its first session being attended by 51 members and the last session by 193 members. Figure 2 shows the increase in the number of voting member states of the UNGA over time.

Figure 2: Number of voting members across time in each UNGA voting session between 1946 and 2022



It must be noted that the UN Digital Library only includes the resolutions that ended up being adopted by the General Assembly. Therefore, as can be seen from the distribution of "Yes" votes in Figure 3 (right), the “Yes” votes are, in general, over-represented in the dataset. The bars to the left of the 0.5 ratio in Figure 3 (right) are explained by the simple majority voting system of UNGA – the number of “Yes” votes only need to be higher than those of “No” votes and abstentions are disregarded.

Figure 3: Distribution of resolution lengths (left) and ratio of “Yes” votes (right)



The length of the resolutions varies widely. With some resolutions containing only a single paragraph and others going on for tens of pages. The distribution of token counts per resolution resembles a lognormal distribution with a mean of around 900 tokens. As can be seen on Figure 3 (left), there are a few resolutions that have more than 10,000 tokens. Token counts are calculated after the removal of stop words from the documents.

III. Methods

To observe the changes in UNGA alliance structures over time since the foundation of the UN, I split the dataset into three periods (Macon, Mucha, & Porter, 2012):

- Early Cold War Period (1946-1960)
- Transitional Period (1961-1990)
- Post-Cold War Period (1991-2022)

This split also allows us to see what countries are added and removed from the network over time as well as the structural changes in the edges connecting the nodes.

A. Dimensionality Reduction with t-SNE

The voting data for each era listed above is represented as three matrices $\mathbf{V}_{N\times R}$ where N is the number of member countries that have voted on at least half of the sessions during that period and R is the number of resolutions that were voted during that period. A value within the matrix $v_{i,j}$ equals 1 if the country i voted "Yes" on the resolution j, -1 if the vote was "No", and 0 otherwise (e.g., abstentions).

To get higher-level country clusters and visualize them on a two-dimensional space, I apply the t-SNE (t-Distributed Stochastic Neighbor Embedding) algorithm on each of the matrices. t-SNE is a non-parametric dimensionality reduction technique that converts similarities between data points to joint probabilities and tries to minimize the Kullback-Leibler divergence between the joint probabilities of the low-dimensional embedding and the high-dimensional data (van der Maaten & Hinton, 2008). It is a variation of the Stochastic Neighbor Embedding, but much easier to optimize and known to produce better visualization "by reducing the tendency to crowd points together in the center of the map."

B. Agglomerative Single-Linkage Clustering

After obtaining the embedded matrices of size $N \times 2$, I apply agglomerative clustering on them with single-linkage. Single-linkage clustering works bottom-up—it starts from individual data points as individual clusters and, at each step, combines two clusters that contain the closest pair of elements not yet belonging to the same cluster as each other. Mathematically, the algorithm decides which clusters to merge at each step by calculating the linkage function (Euclidian distance) for each cluster pair.

$$D(X, Y) = \min_{x \in X, y \in Y} \sqrt{\sum_{i=1}^{2} (x_i - y_i)^2}$$

Where X and Y are two clusters at any point in the clustering process until all data points are merged into a single cluster. Then I decide on the number of clusters by examining the distance between clusters in dendogram plots as well as each country's location in the embedded space. t-SNE coupled with agglomerative single-linkage clustering yields the most intuitive and clear higher-level country clusters as can be seen in the results section.

C. Text Pre-processing and Topic Modeling

I pre-process each resolution document in the corpus first by removing the "stop words" that would not add any useful information to the topic model. These words are the most common words in the English language such as "is", "are", "in", or "on". This process allows the topic model to run more efficiently and capture the most relevant information in the documents. After splitting each document into an array of unigram tokens, I lowercase and lemmatize each token using the NLTK WordNet lemmatizer. An example document of two sentences such as `"United Nations recommendation to member states. Nuclear non-proliferation."` would be processed to show `["unite", "nation", "recommendation", "member", "state", "nuclear", "non-proliferation"].`

After pre-processing, I extract policy areas from the textual content of each document through a Latent Dirichlet Allocation topic model (LDA). A topic model such as LDA views each document as a "bag of words", a collection of words without order or structure. It aims to discover a latent distribution of "topics"— probabilities of certain terms appearing in a document together (Beli, Ng, & Jordan, 2003). The topic model proposes a data generating process in which the probability that a given document, d, contains a given term, w, is the sum of products of conditional probabilities for each of the topics, t, to include that term, where N in this case is the number of topics.

$$\mathrm{P}(W = w \mid D = d) = \sum_{t=1}^{N} \mathrm{P}(W = w \mid T = t)\, \mathrm{P}(T = t \mid D = d)$$

For instance, the words "disarmament", "nuclear", "treaty", and "weapon" could constitute the top five most salient terms of a topic that we may label as "Nuclear Disarmament." Each resolution document has probabilities associated with each topic, and each topic has probabilities associated with each term. Perplexity and topic coherence are among the metrics that can be used to determine the number of topics to be extracted from the corpus. A lower perplexity score and a higher coherence score are desired. I went with a 14 topics based on these metrics and ease of interpretation.

After training the LDA model, I predict topic distributions for each resolution document. Constructing a document-topic matrix $\mathbf{M}_{R\times T}$ where R is the number of resolutions and T is the number of topics. Where $m_{i,j}$ is the weight of the topic j on the resolution i. And $sum(\boldsymbol{m}_i) = 1$ for each i.

D. Social Network Analysis

To represent alliances within each time period and international policy area with a graph data structure, I construct 42 (3 eras × 14 policy areas) adjacency matrices $\mathbf{A}^t_{N\times N}$ where N is the number of UN member countries that have voted in at least half of the resolutions passed during the given period, and t is a policy area. Each entry in the adjacency matrices can be defined as the following.

$$a^t_{i,j} = \left(\left(\boldsymbol{v}_i \circ \boldsymbol{v}_j\right) \times \mathbf{M} \right)_t$$

Where $\boldsymbol{v}_i$ is the row vector of the vote matrix $\mathbf{V}_{N\times R}$ that corresponds to the votes of the country i for all resolutions R, and $\mathbf{M}$ is the document-topic matrix described above. In other words, each entry in the adjacency matrices gives us the "strength of alliance" between two countries for a given issue area (topic). If a document belongs to the topic of "nuclear disarmament" and two countries casted the same vote, their "strength of alliance" is increased in the "nuclear disarmament” issue area, proportional to the topic’s weight in the given document. The opposite, is true for countries that have voted against each other – their “strength of alliance” is decreased. Therefore, if two countries voted against each other consistently on a given topic, their entry in the relevant adjacency matrix is going to be lower than two countries that voted in unison most of the time. Later, these forty-two alliance matrices are used to build graphs where each node is a country and the edges between the nodes are determined by the entries in the alliance matrices. For visualization of these network graphs, the Fruchterman-Reingold algorithm (spring layout or force-directed graph drawing algorithm) was leveraged to represent nodes that have more weights in their edges (links) closer to each other and nodes that do not share any links farther away in the graph (Fruchterman & Reingold, 1991). Individual countries make up the nodes in the graph and the weight of the edges between the nodes are determined by the “strength of alliance” metric characterizing the relationship of the nodes. The network visualizations generated by this algorithm and setup, therefore, give us the chance to visually discriminate the defining differences in alliance structures across issue areas and time periods.

IV. Results

A. Higher-Level Country Clusters

Using the t-SNE algorithm with PCA initialization and the perplexity parameters of [5, 5, 40], I obtain three embedded two-dimensional spaces of the vote matrix for each of the time periods specified above. As can be seen in figures 4 through 9, t-SNE reveals arguably clear-cut clusters for each of the periods and member countries seem to be very well separated from each other within these embedded spaces based on their voting patterns across all issue areas discussed in the UNGA. t-SNE is non-deterministic, however, initialization of the model with different random seeds as well as varying perplexity parameters did not show any significant divergence from the figures on display, reinforcing the robustness and consistency of the results. After t-SNE, the assignment of countries to a cluster appears to be a simple task that can be done visually. After experimenting with some clustering algorithms like K-means, complete-linkage, and average-linkage, I chose to go with single-linkage clustering as it yielded the most intuitive results taking into account the scatterplots over the lower-dimensional space.

Figure 4: Two-dimensional t-SNE projection of countries during the Early Cold War Period

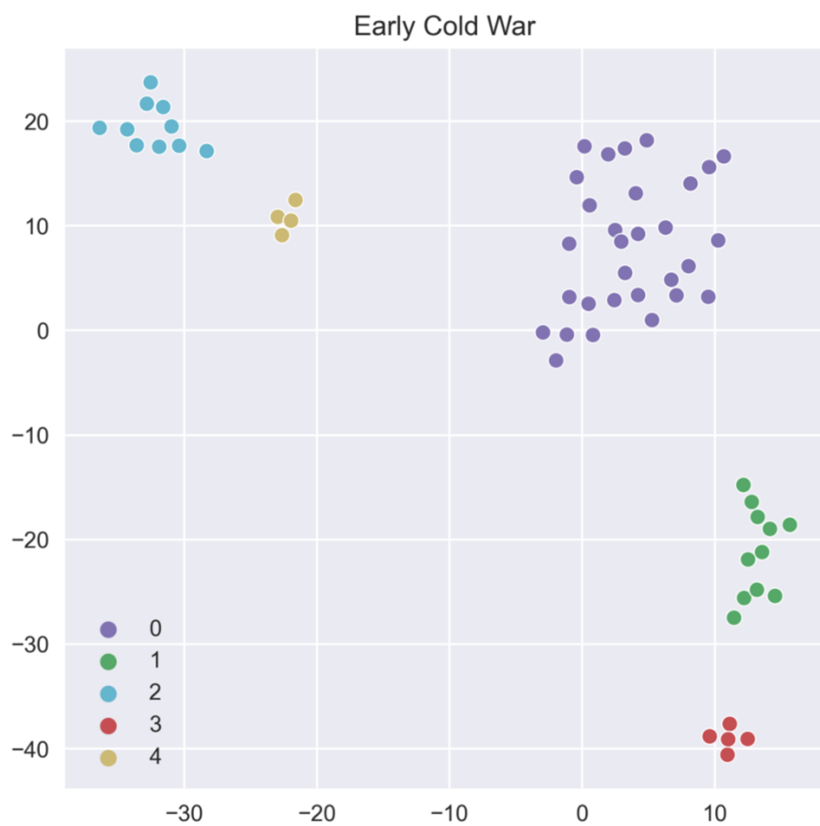


Figure 5: Hierarchical Clustering Dendogram for the Early Cold War Period with clusters corresponding to colors

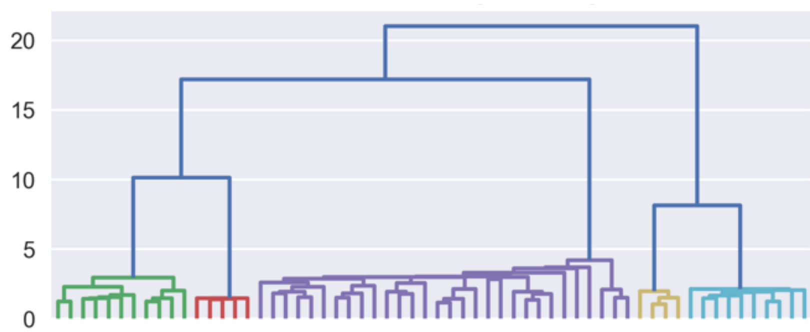


Table 1: Clustering output for the Early Cold War Period

Period	Cluster	Country List
Early Cold War (1946-1960)	0	ARGENTINA, BOLIVIA, BRAZIL, CHILE, CHINA, COLOMBIA, COSTA RICA, CUBA, DOMINICAN REPUBLIC, ECUADOR, EL SALVADOR, ETHIOPIA, GREECE, GUATEMALA, HAITI, HONDURAS, IRAN, ISRAEL, LIBERIA, MEXICO, NICARAGUA, PAKISTAN, PANAMA, PARAGUAY, PERU, PHILIPPINES, THAILAND, TURKEY, URUGUAY, VENEZUELA
	1	AFGHANISTAN, EGYPT, INDIA, INDONESIA, IRAQ, LEBANON, MYANMAR, SAUDI ARABIA, SYRIA, YEMEN, YUGOSLAVIA
	2	AUSTRALIA, BELGIUM, CANADA, FRANCE, LUXEMBOURG, NETHERLANDS, NEW ZEALAND, SOUTH AFRICA, UNITED KINGDOM, UNITED STATES
	3	BELARUS, CZECHOSLOVAKIA, POLAND, UKRAINE, USSR
	4	DENMARK, ICELAND, NORWAY, SWEDEN

The higher-level country clusters across the three periods reveal some interesting characteristics. The Early Cold War era, expectedly, reveals the polarized nature of international politics during the 1946-1960 period. Western bloc, or the first world, are situated in the upper-left quadrant of the scatterplot. One thing to notice is the relative distance of Scandinavian nations (cluster 4) from the rest of the West (cluster 2). Also, we can easily notice the ideological divide of the period—USSR and the countries under the Soviet sphere of influence are situated at the lower-right quadrant of the plot (cluster 3). The other two clusters (0 and 1) largely consist of the members of the non-aligned movement led by Yugoslavia. Though it is still possible to distinguish which members of the non-aligned movement were relatively ideologically more adjacent to which camp.

The so-called Third World seems to be split into two. Cluster 0 is mostly made up of Latin American members in addition to the ideologically even more West-leaning states of the time, such as Israel and Turkey. Whereas the Cluster 1 is made up of the ideologically East-leaning states of the non-aligned movement such as Yugoslavia and India, as well as some of the Muslim states of the time: Afghanistan, Egypt, Iraq, and Yemen. It should be noted that the pre-revolution Iran seems to be one of the few Muslim states that is assigned to the slightly more West-leaning cluster of the non-aligned movement. This period also contains the lowest number of countries (n = 60**)** as the wave of colonization had not started then.

Figure 6: Two-dimensional t-SNE projection of countries during the Transitional Period

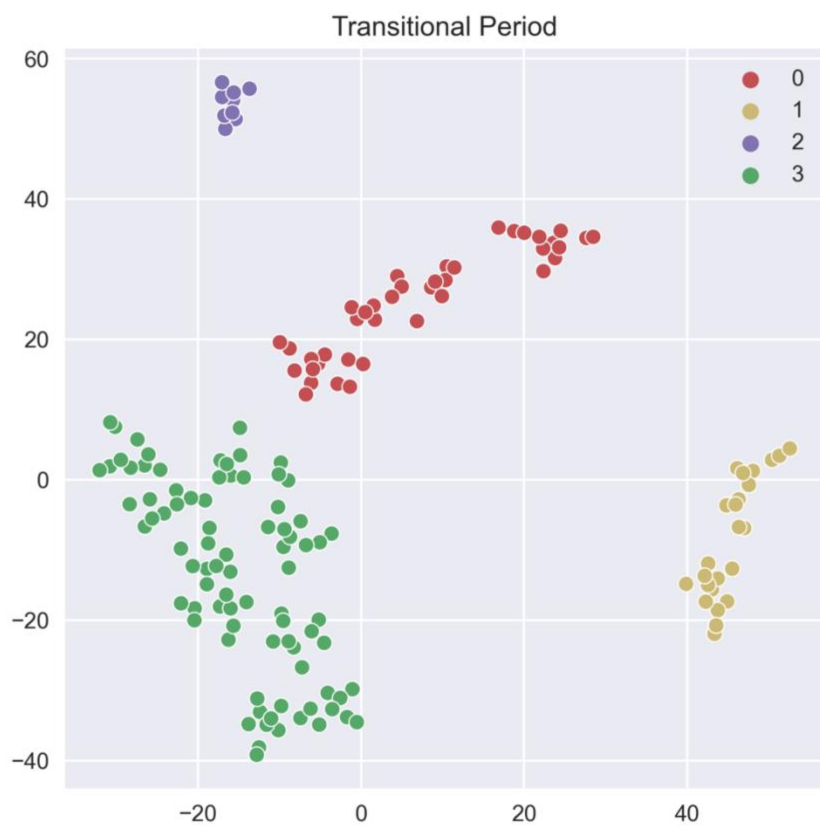


Figure 7: Hierarchical Clustering Dendogram for the Transitional Period with clusters corresponding to colors

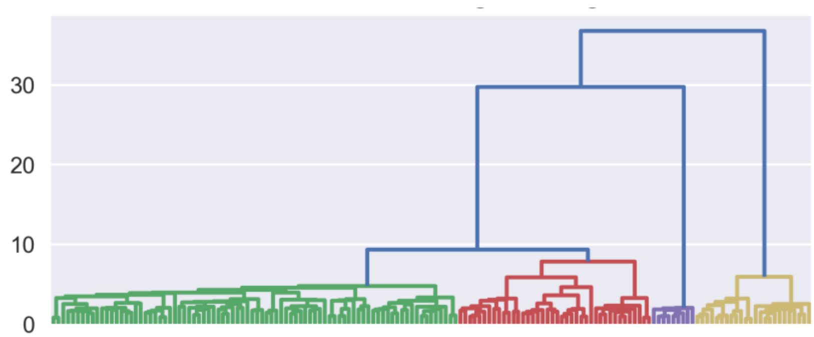


Table 2: Clustering output for the Transitional Period

Period	Cluster	Country List
Trainsitional Period (1961-1990)	0	ALBANIA, ANGOLA, ANTIGUA AND BARBUDA, BAHAMAS, BAHRAIN, BANGLADESH, BELIZE, BHUTAN, BOTSWANA, CAPE VERDE, COMOROS, DJIBOUTI, DOMINICA, EQUATORIAL GUINEA, FIJI, GAMBIA, GRENADA, GUINEA-BISSAU, LESOTHO, MALAWI, MALDIVES, MALTA, MAURITIUS, MOZAMBIQUE, OMAN, PAPUA NEW GUINEA, QATAR, SAINT LUCIA, SAINT VINCENT AND THE GRENADINES, SAMOA, SAO TOME AND PRINCIPE, SEYCHELLES, SOLOMON ISLANDS, SOUTH AFRICA, SURINAM, SWAZILAND, UNITED ARAB EMIRATES, VANUATU, VIET NAM, ZIMBABWE
	1	AUSTRALIA, AUSTRIA, BELGIUM, CANADA, DENMARK, FINLAND, FRANCE, GERMANY, FEDERAL REPUBLIC OF, GREECE, ICELAND, IRELAND, ISRAEL, ITALY, JAPAN, LUXEMBOURG, NETHERLANDS, NEW ZEALAND, NORWAY, PORTUGAL, SPAIN, SWEDEN, TURKEY, UNITED KINGDOM, UNITED STATES
	2	BELARUS, BULGARIA, CZECHOSLOVAKIA, GERMAN DEMOCRATIC REPUBLIC, HUNGARY, MONGOLIA, POLAND, UKRAINE, USSR
	3	AFGHANISTAN, ALGERIA, ARGENTINA, BARBADOS, BENIN, BOLIVIA, BRAZIL, BURKINA FASO, BURUNDI, CAMBODIA, CAMEROON, CENTRAL AFRICAN REPUBLIC, CHAD, CHILE, CHINA, COLOMBIA, CONGO, COSTA RICA, CUBA, CYPRUS, DEMOCRATIC REPUBLIC OF THE CONGO, DEMOCRATIC YEMEN, DOMINICAN REPUBLIC, ECUADOR, EGYPT, EL SALVADOR, ETHIOPIA, GABON, GHANA, GUATEMALA, GUINEA, GUYANA, HAITI, HONDURAS, INDIA, INDONESIA, IRAN, IRAQ, IVORY COAST, JAMAICA, JORDAN, KENYA, KUWAIT, LAOS, LEBANON, LIBERIA, LIBYA, MADAGASCAR, MALAYSIA, MALI, MAURITANIA, MEXICO, MOROCCO, MYANMAR, NEPAL, NICARAGUA, NIGER, NIGERIA, PAKISTAN, PANAMA, PARAGUAY, PERU, PHILIPPINES, ROMANIA, RWANDA, SAUDI ARABIA, SENEGAL, SIERRA LEONE, SINGAPORE, SOMALIA, SRI LANKA, SUDAN, SYRIA, THAILAND, TOGO, TRINIDAD AND TOBAGO, TUNISIA, UGANDA, UNITED REPUBLIC OF TANZANIA, URUGUAY, VENEZUELA, YEMEN, YUGOSLAVIA, ZAMBIA

As we move on to the Transitional Period (1961-1990), the most striking difference from the Early Cold War Period is the threefold increase in the number of member countries who have attended at least half of the voting sessions of the UNGA. This can be attributed to the wave of decolonization, and the integration of the former colonies to the international political system through multilateral intergovernmental organizations such as the UN. During the transitional period, it can be noticed that the ideological split is more consolidated, with the Scandinavian countries, as well as Turkey and Israel grouped in the Global West cluster (cluster 1) reflecting the US strategy prevalent during the period. Number of countries in the Soviet sphere of influence also increased with the addition of Bulgaria, East Germany, Hungary, and Mongolia (cluster 2). The other two clusters (0 and 3) contain the bulk of the new members and the growing non-aligned movement analogous to the clusters 0 and 1 of the Early Cold War Period. Most of the less developed nations of Africa and Asia as well as the small island developing states across the Atlantic and Pacific are split between these two clusters although the Latin American members remained as a group (cluster 3). Albania, Romania, Cyprus, and Yugoslavia are the only European countries not assigned to the ideological West or East.

Figure 8: Two-dimensional t-SNE projection of countries during the Post-Cold War Period

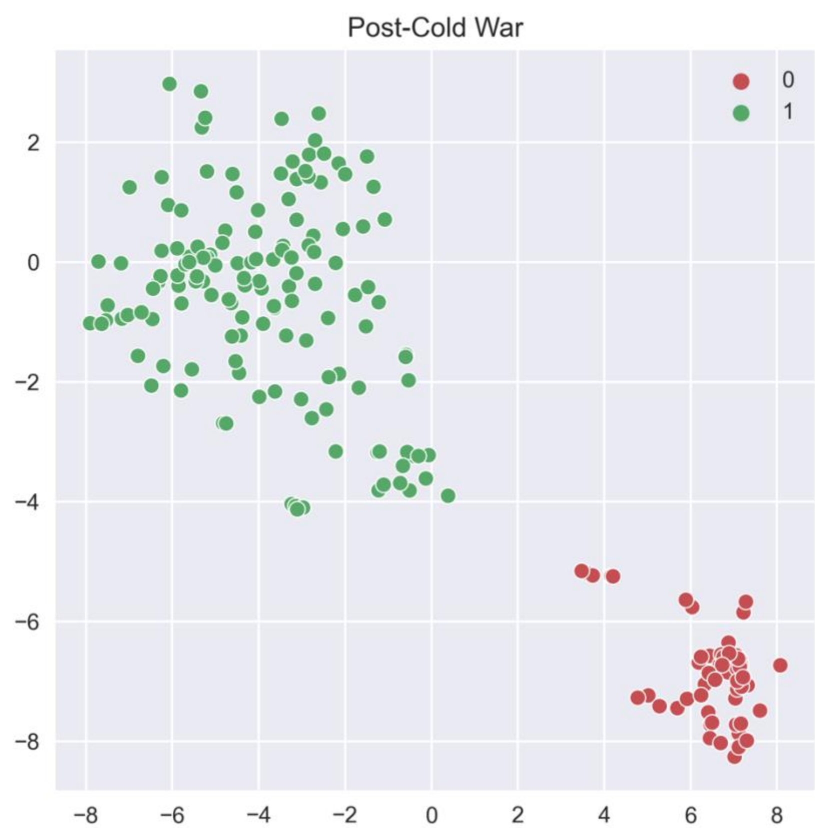


Figure 9: Hierarchical Clustering Dendogram for the Post-Cold War Period with clusters corresponding to colors

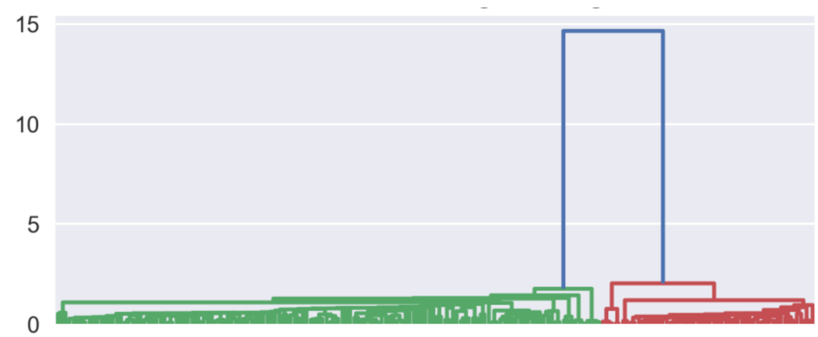


Table 3: Clustering output for the Post-Cold War Period

Period	Cluster	Country List
Post-Cold War (1991-2022)	0	ALBANIA, ANDORRA, AUSTRALIA, AUSTRIA, BELGIUM, BOSNIA AND HERZEGOVINA, BULGARIA, CANADA, CROATIA, CYPRUS, CZECHIA, DENMARK, ESTONIA, FINLAND, FRANCE, GEORGIA, GERMANY, GREECE, HUNGARY, ICELAND, IRELAND, ISRAEL, ITALY, JAPAN, LATVIA, LIECHTENSTEIN, LITHUANIA, LUXEMBOURG, MALTA, MARSHALL ISLANDS, MICRONESIA (FEDERATED STATES OF), MONACO, MONTENEGRO, NETHERLANDS, NEW ZEALAND, NORTH MACEDONIA, NORWAY, PALAU, POLAND, PORTUGAL, REPUBLIC OF KOREA, REPUBLIC OF MOLDOVA, ROMANIA, SAN MARINO, SERBIA, SLOVAKIA, SLOVENIA, SPAIN, SWEDEN, SWITZERLAND, TURKEY, UKRAINE, UNITED KINGDOM, UNITED STATES
	1	AFGHANISTAN, ALGERIA, ANGOLA, ANTIGUA AND BARBUDA, ARGENTINA, ARMENIA, AZERBAIJAN, BAHAMAS, BAHRAIN, BANGLADESH, BARBADOS, BELARUS, BELIZE, BENIN, BHUTAN, BOLIVIA, BOTSWANA, BRAZIL, BRUNEI DARUSSALAM, BURKINA FASO, BURUNDI, CAMBODIA, CAMEROON, CAPE VERDE, CENTRAL AFRICAN REPUBLIC, CHAD, CHILE, CHINA, COLOMBIA, COMOROS, CONGO, COSTA RICA, CUBA, DEMOCRATIC PEOPLE'S REPUBLIC OF KOREA, DEMOCRATIC REPUBLIC OF THE CONGO, DJIBOUTI, DOMINICA, DOMINICAN REPUBLIC, ECUADOR, EGYPT, EL SALVADOR, EQUATORIAL GUINEA, ERITREA, ETHIOPIA, FIJI, GABON, GAMBIA, GHANA, GRENADA, GUATEMALA, GUINEA, GUINEA-BISSAU, GUYANA, HAITI, HONDURAS, INDIA, INDONESIA, IRAN, IRAQ, IVORY COAST, JAMAICA, JORDAN, KAZAKHSTAN, KENYA, KIRIBATI, KUWAIT, KYRGYZSTAN, LAOS, LEBANON, LESOTHO, LIBERIA, LIBYA, MADAGASCAR, MALAWI, MALAYSIA, MALDIVES, MALI, MAURITANIA, MAURITIUS, MEXICO, MONGOLIA, MOROCCO, MOZAMBIQUE, MYANMAR, NAMIBIA, NAURU, NEPAL, NICARAGUA, NIGER, NIGERIA, OMAN, PAKISTAN, PANAMA, PAPUA NEW GUINEA, PARAGUAY, PERU, PHILIPPINES, QATAR, RUSSIAN FEDERATION, RWANDA, SAINT KITTS AND NEVIS, SAINT LUCIA, SAINT VINCENT AND THE GRENADINES, SAMOA, SAO TOME AND PRINCIPE, SAUDI ARABIA, SENEGAL, SEYCHELLES, SIERRA LEONE, SINGAPORE, SOLOMON ISLANDS, SOMALIA, SOUTH AFRICA, SRI LANKA, SUDAN, SURINAM, SWAZILAND, SYRIA, TAJIKISTAN, THAILAND, TIMOR-LESTE, TOGO, TONGA, TRINIDAD AND TOBAGO, TUNISIA, TURKMENISTAN, TUVALU, UGANDA, UNITED ARAB EMIRATES, UNITED REPUBLIC OF TANZANIA, URUGUAY, UZBEKISTAN, VANUATU, VENEZUELA, VIET NAM, YEMEN, ZAMBIA, ZIMBABWE

Compared to the Cold War, The Post-Cold War (1991-2022) period have been argued to be defined by less ideological divide and more cultural divide along the lines of world's major civilizations

(Huntington, 2000). That divide, if it existed in the UNGA voting patterns, was not revealed by the manifold that was learned by the t-SNE algorithm. Instead, we see two camps that are more loosely defined yet still make sense within the ideological and geopolitical conjuncture of the time. With the collapse of the Berlin wall in 1989 and the dissolution of the USSR in 1991 and Yugoslavia in 1992, the newly emerging countries have been admitted to the UN and have found their ideological blocs. Former communist/socialist states of Eastern Europe such as Poland, Ukraine, Hungary, and Bulgaria, have transitioned to the Global North (cluster 0) following the end of the Cold War. Whereas the Turkic republics that were formerly a part of USSR, such as Turkmenistan and Uzbekistan, have maintained their adjacency to Russia. The higher-level picture we see in UNGA voting patterns during this period seem to be most aligned with the framework defined by the Global North vs. Global South quarrel. Though as I will show with the Social Network Analysis, we see convincing evidence against this framework when we examine the characteristics of interstate alliances on a case-by-case basis by taking the changing incentives surrounding various policy areas into account.

B. UNGA Issue Areas

In order to extract the prevalent issue areas that were discussed on the floor of the UNGA since its foundation until March 2022, I trained eight topic models using the complete corpus with varying number of topics. Although none are perfect, there are a few metrics to evaluate the usefulness of a given topic when manual inspection is infeasible due to the sheer number of documents in a corpus or lengthy texts (e.g., books, articles). One of such metrics is perplexity (also sometimes referred as the held out log-likelihood), which is a measure of how successfully a trained topic model predicts the held out topic distribution on unseen documents. Although this paper's goal is not to make predictions and despite plausible criticisms against the metric's ability to measure semantic meaningfulness (Chang, Boyd-Graber, Wang, Gerrish, & Blei, 2009), I show them along with the coherence scores in figure 10. Whereas perplexity does not capture context (relationship between words in a topic or topics in a document), coherence aims to measure the conditional likelihood of the co-occurrence of words in a topic. High coherence score with low perplexity is desired in a topic model. Figure 10 shows that when topic number equals fourteen, the coherence score rises dramatically and reduces afterwards. Also, perplexity scores start to decline at a slightly lower rate after fourteen topics. So, after examining a sample of the documents manually, I chose to go with the model with fourteen topics that correspond to fourteen issue areas discussed at the UNGA which will give me issue specific insights into how international cooperation and conflict evolved over time.

Figure 10: Metrics for topic models with varying number of topics

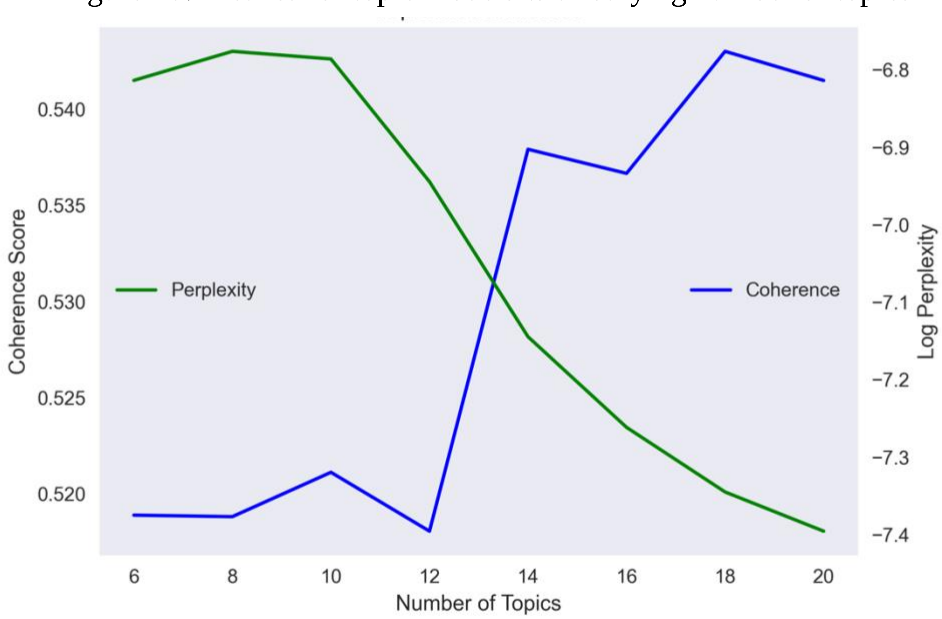


Table 4: Extracted topics and ten of the top twenty most relevant terms, labeled with corresponding issue areas

Label	Term	Relevance	Label	Term	Relevance	Label	Term	Relevance	Label	Term	Relevance
Human Rights	right	-0.889	UN Appointments and Administration	republic	-1.625	Outer Space, Science, and Technology	information	-1.448	Sustainable Development and Poverty Eradication	development	-1.796
	human	-1.076		board	-1.856		space	-1.778		sustainable	-2.375
	freedom	-2.252		membership	-2.021		outer	-2.069		food	-2.695
	cultural	-2.501		year	-2.097		department	-2.079		support	-2.962
	promotion	-2.529		pension	-2.165		scientific	-2.500		global	-2.995
	fundamental	-2.539		administrative	-2.177		activity	-2.530		world	-3.013
	declaration	-2.693		auditor	-2.285		secretariat	-2.533		need	-3.072
	covenant	-2.744		appointment	-2.416		language	-2.575		poverty	-3.118
	protection	-2.818		fund	-2.420		technology	-2.578		goal	-3.121
	respect	-2.818		appoint	-2.517		exploration	-2.916		disaster	-3.128
Discrimination and Violence Against Women and Children	woman	-1.477	Decolonization and Self Determination	territory	-1.311	Fight Against Global Crime	crime	-1.870	Oceans and Law of the Sea	convention	-1.208
	child	-1.583		administer	-1.778		drug	-2.095		law	-1.456
	discrimination	-2.252		power	-1.800		traffic	-2.172		party	-1.859
	education	-2.302		island	-2.231		criminal	-2.264		sea	-2.069
	violence	-2.474		people	-2.254		prevention	-2.327		marine	-2.209
	girl	-2.494		territorial	-2.363		cooperation	-2.333		ocean	-2.307
	person	-2.501		government	-2.472		illicit	-2.533		torture	-2.502
	youth	-2.510		guam	-2.599		justice	-2.574		protocol	-2.508
	health	-2.561		caribbean	-2.730		terrorism	-2.611		maritime	-2.606
	racial	-2.586		independence	-2.822		prevent	-2.703		fish	-2.736
Nuclear Disarmament	disarmament	-1.610	UN Membership	shall	-1.836	South Africa and Apartheid	africa	-1.439			
	nuclear	-1.638		article	-2.219		south	-1.587			
	weapon	-1.787		commission	-2.452		african	-1.995			
	treaty	-2.174		may	-2.536		apartheid	-2.040			
	arm	-2.409		charter	-2.690		people	-2.106			
	conference	-2.551		present	-2.873		namibia	-2.210			
	security	-2.627		work	-2.909		independence	-2.372			
	use	-2.827		paragraph	-2.931		colonial	-2.421			
	peace	-2.837		draft	-2.974		regime	-2.604			
	negotiation	-2.936		council	-2.996		southern	-2.645			
Financing of UN Missions	dollar	-1.850	Arab-Israeli Conflict and Refugees	refugee	-1.798	Industrial Development	economic	-1.954			
	mission	-2.053		humanitarian	-1.926		programme	-2.002			
	budget	-2.099		assistance	-2.199		development	-2.112			
	amount	-2.145		occupy	-2.367		organization	-2.174			
	june	-2.197		palestine	-2.489		conference	-2.410			
	period	-2.223		afghanistan	-2.620		social	-2.411			
	staff	-2.231		relief	-2.680		country	-2.530			
	assessment	-2.280		displace	-2.695		council	-2.687			
	estimate	-2.401		arab	-2.735		assistance	-2.818			
	finance	-2.427		israel	-2.778		trade	-2.893			

Table 4 shows the most relevant terms for each of the fourteen topics I will use as issue areas in the following section. Although there are certain topics that overlap, we can still distinguish all topics from the others. Terms are ranked by their “relevance” to the given topic (Sievert & Shirley, 2014).

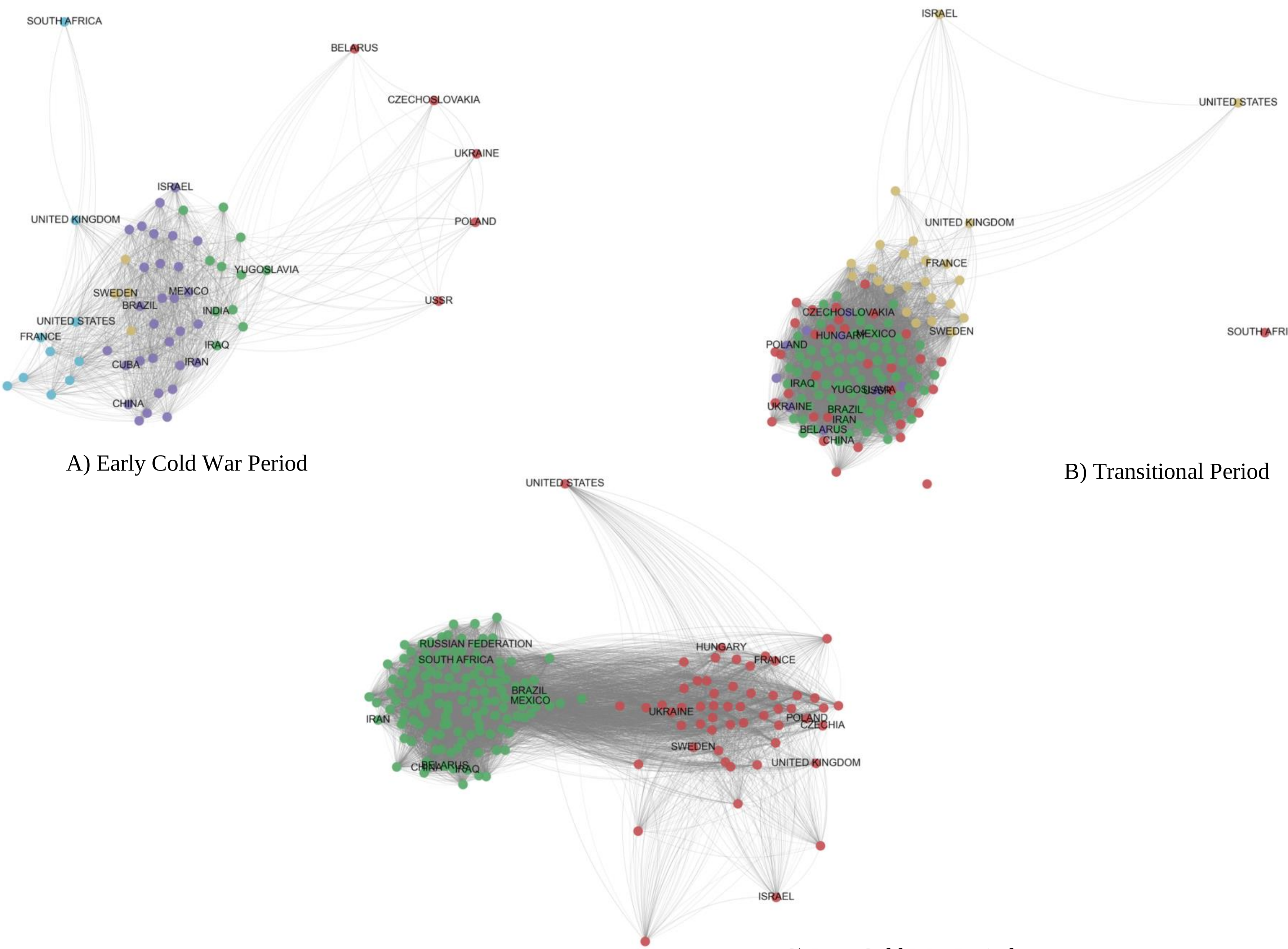
C. Issue Specific Alliance Structures

Using the topic model described in the above section, I calculate document-topic distributions for each resolution in the corpus. This allows me to represent a document as a size 14 vector where each element of the vector is the weight of an issue area for that document. These vectors are then used to compute an alliance metric per time period and issue area, which are represented as social networks that can give us insights on specific voting patterns.

Example 1: Human Rights

Human rights have been the subject of many UNGA resolutions since its founding. Examining the alliance networks across the three time periods reveals some interesting dynamics. In the Early Cold War period (figure 11, A), we can see a relative isolation of the Warsaw Pact members (in red) with Yugoslavia at the closer edge of the large central community. Expectedly, most members of the western capitalist bloc (in blue) situated at the opposite end at the end of the Fruchterman-Reingold simulation. Moreover, the Scandinavian countries (in yellow) seem to be more closely aligned with non-aligned movement than the western bloc in this issue area.

Figure 11: Network visualizations for the Human Rights issue area



A) Early Cold War Period

B) Transitional Period

C) Post-Cold War Period

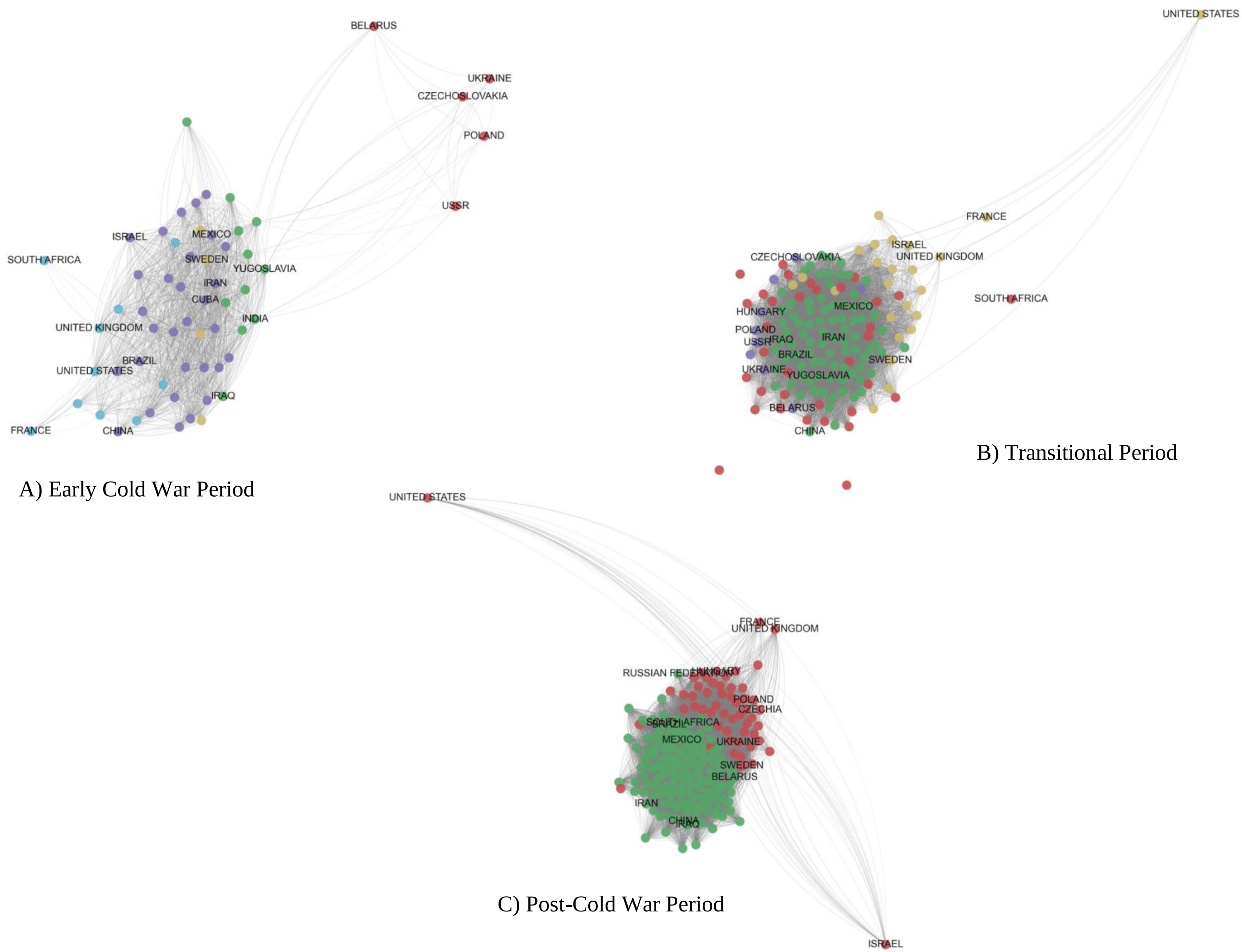
During the Transitional Period, the alliance structure in the Human Rights issue area have changed significantly. The Western bloc (in yellow), especially Israel and the United States now seem to be the most isolated of the cohort. Whereas, red, yellow, and purple nodes are tightly grouped together, indicating no significant disagreements between the Soviet bloc and the non-aligned movement unlike the early years of the Cold War. South Africa, possibly due to apartheid politics, seem to be isolated from the rest of the world throughout the Cold War.

The Post-Cold War Period, on the other hand, resembles the higher-level country clusters discussed above, with a clear divide between the Global North (in red) and the Global South (in green). Ex-communist countries like Hungary and Ukraine have started to align more with the Global North in the Human Rights issue area with the end of the Cold War. South Africa, with the release of Nelson Mandela from prison in 1990 and subsequent legal and political developments of early 1990s that undermined the remnants of the apartheid politics, started to align more with the Global South concerning human rights related topics in the UNGA.

Example 2: Nuclear Disarmament

The picture of alliance structures concerning Nuclear Disarmament during the early years of the Cold War somewhat resembles the Human Rights issue area voting patterns of the same period. Soviet bloc (in red) differentiated from the rest of the world with France at the other end of the network. Yellow nodes representing the Scandinavian countries as well as some of the blue nodes of the "free world" (Canada and New Zealand) are now closer to the Third World than to their western allies, an interesting deviation away from the higher-level country clusters.

Figure 12: Network visualizations for the Nuclear Disarmament issue area



A) Early Cold War Period

B) Transitional Period

C) Post-Cold War Period

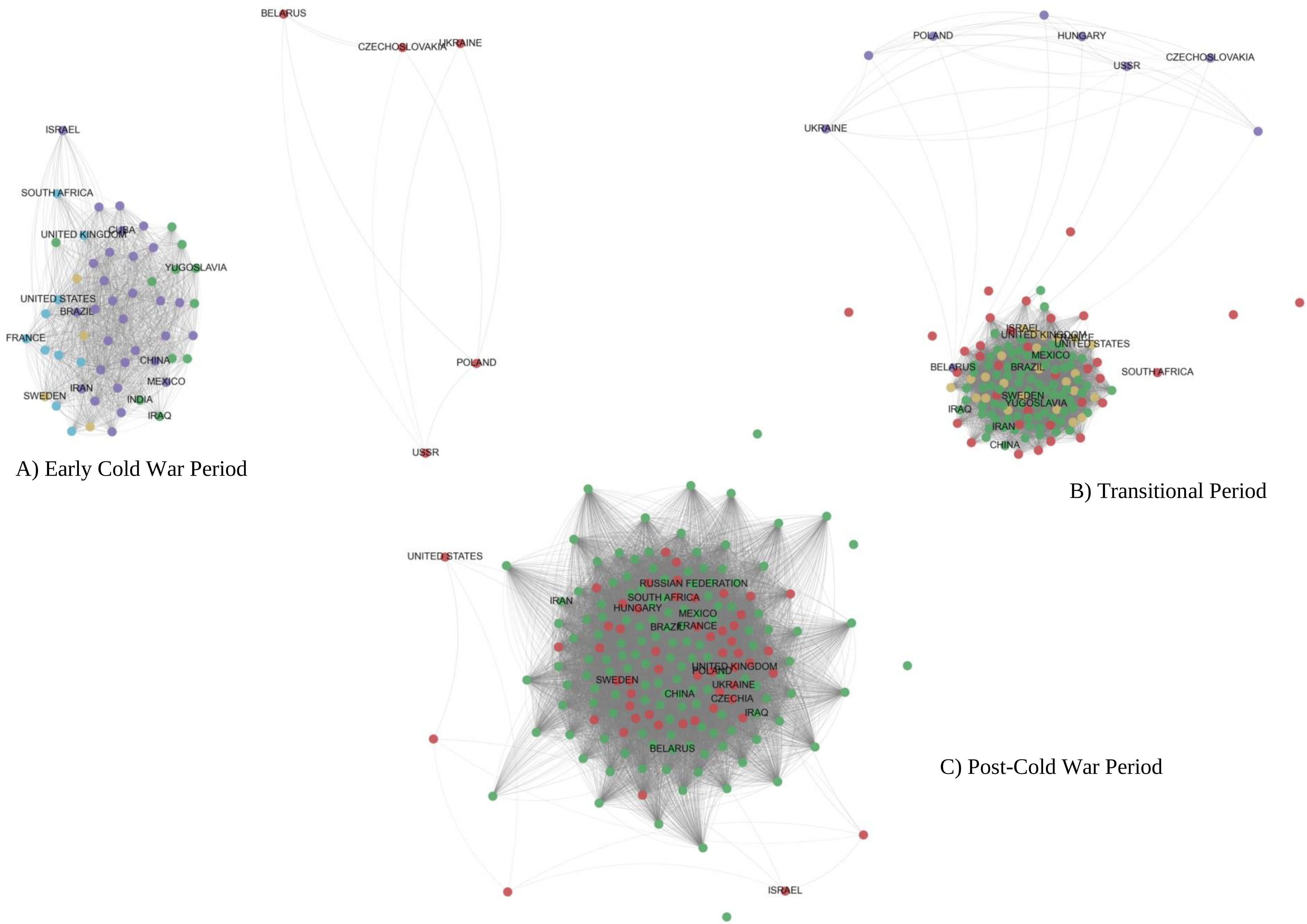
During the Transitional Period of the Cold War, we see a clear isolation of the United States in the Nuclear Disarmament issue area. And the newly independent countries have formed a coalition with the Eastern bloc. Finland, Austria, Ireland, and Iceland have remained relatively from the rest of the West when it came to nuclear weapons (yellow nodes in the central community).

The Post-Cold War period nuclear disarmament discussions also show some deviations from the mean. The most prominent one is Russia, now closer than ever to the Western World and other nuclear powers like France, and the UK, illustrating the evolving nature of nuclear diplomacy after the Cold War. Also interestingly, Palau and Marshall Islands are the two red nodes in the green cluster. Those two island states, as signatories of the international agreement Compact of Free Association (COFA), benefit from access to a range of economic and military provisions provided by the US. Possibly due to this partnership these states, along with Micronesia, voted almost in perfect synchronization with the US. However, the Nuclear Disarmament issue area seems to be an exception that can easily be captured by the network representations. Possibly due to polarizing talks around the Iran nuclear deal (which was also backed by Russian Foreign Minister Sergey Lavrov), Iran remained at the very far end of the network vis-à-vis the Global North. Israel and the US are again separated from the rest.

Example 3: Financing of UN Missions

A big bulk of UN's resources are spent to fund employee salaries and pensions in the permanent UN missions in member countries. This is an issue area of the UNGA where one would expect diplomats to agree regardless of ideological differences, at least relative to other issue areas.

Figure 13: Network visualizations for the Financing of UN Missions



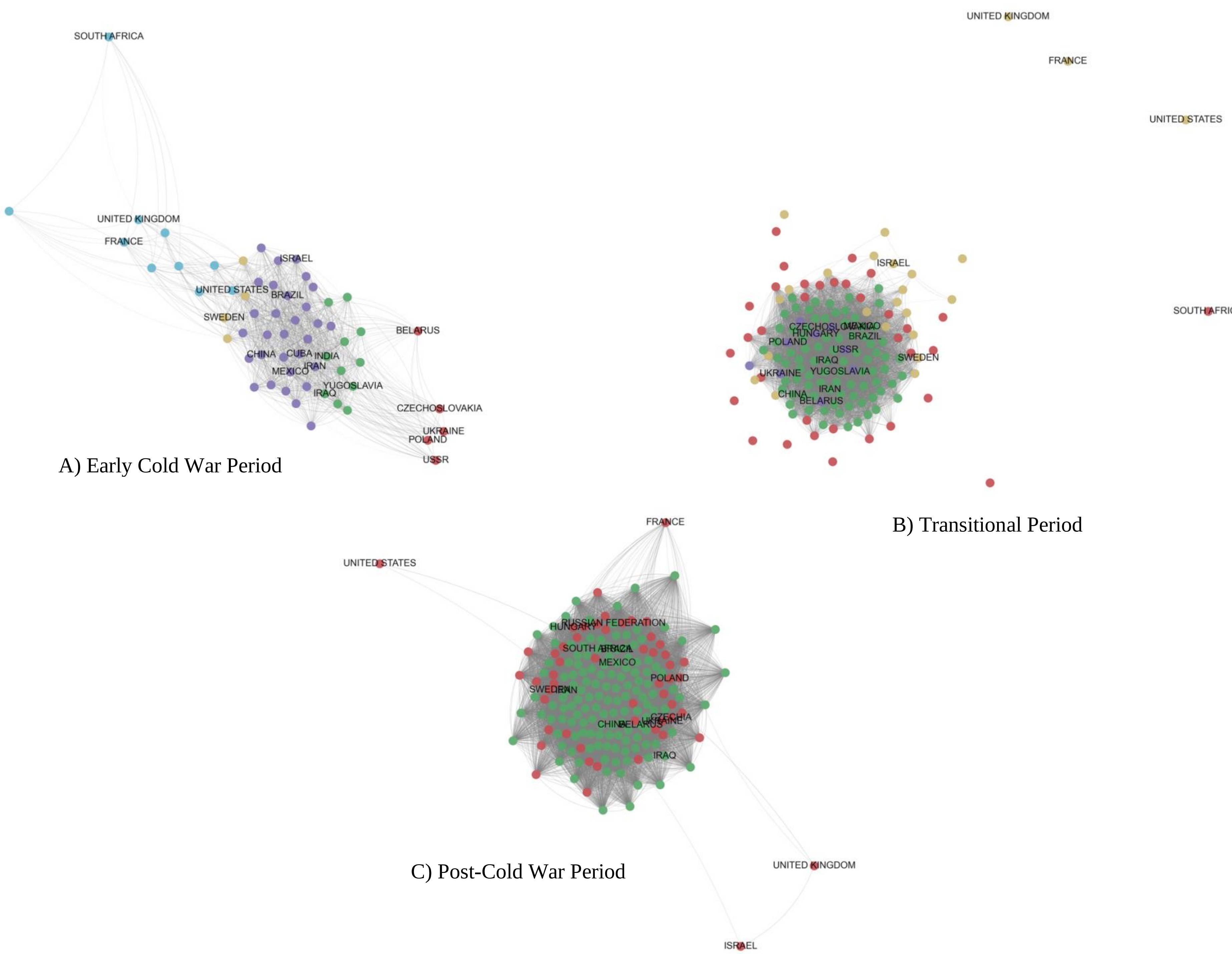
Nonetheless, the we see an isolation of the Warsaw Pact members throughout the Cold War with the rest of the world, especially in the Transition Period getting along without displaying the divergences we in the higher-level country clusters. After the end of the Cold War, it can be said that the funding of the UN permanent missions has become less ideologically divided. The North-South divide we see in the Post-Cold War higher-level clusters are significantly blurred in mostly administrative matters. This Post-Cold War pattern is sustained in “UN Appointments and Administration” and “UN Membership” issue areas as well.

Example 4: Decolonization and Self Determination

Decolonization and Self Determination issue area is the only issue that has been examined so far that show no clear isolation of the Soviet bloc in the Early Cold War period. Instead, United Kingdom, Belgium, and France (in blue) remain on the far edge of the network with ties only to few members belonging to the western camp. Whereas, opposite to the patterns examined above, the Warsaw Pact members (in red) remain connected to a bulk of the non-aligned movement of the early Cold War. United States, along with Scandinavian countries, is situated noticeably closer to the non-aligned movement compared to other members of the western bloc with colonial pasts.

This pattern changes during the Transitional Period. The most striking fact is the almost complete isolation of countries with a colonial past such as the UK, France, and Belgium. The wave of decolonization that gained pace during the 1960s led to end of these countries’ rule in mainland Africa by 1977.

Figure 14: Network visualizations for the Decolonization and Self Determination issue area



A) Early Cold War Period

B) Transitional Period

C) Post-Cold War Period

With the change in domestic political rhetoric and shifting ideological landscape surrounding colonial activities, the Global North seems to have integrated into the now relatively more mainstream view of anticolonialism after the end of the Cold War.

V. Discussion and Conclusion

Through the framework and methodologies proposed in this paper, my research makes three main contributions. Firstly, I aim to make an empirical contribution by introducing a previously virtually unexplored corpus of diplomatic text reflecting many interesting dynamics of the international politics throughout the last 80 years. Hopefully, this dataset could lead to a range of novel research in the newly emerging field of Computational International Relations. Secondly, I aim to make a theoretical contribution by reevaluating and challenging some of the long-standing International Relations theories by providing a more adaptive and time-sensitive framework in analyzing international alliances and adversarial relationships across different policy areas. And finally, I aim to make a methodological contribution by utilizing a range of computational approaches on diplomatic text to formulate both overarching and latent patterns in international affairs. Such patterns and networks have only previously been explored in qualitative studies which were generally either limited in scale or high in cost, and possibly more prone to bias. Through transparent processing of publicly accessible data in a reproducible way, I aim to remedy some of these shortcomings of legacy methodologies previously embraced in the field.

The corpus curated for this research can be utilized to perform a range of other tasks as well. Recent advances in deep learning architectures and language models that can process long sequences efficiently, make it possible to engage in classification tasks that could help predict country votes on draft resolution documents, possibly uncovering certain nuances in diplomatic behavior and the factors affecting it. Furthermore, it could be possible to make projections into the future about the UN policy landscape and how it is going to evolve. Resolutions discussed on the floor of UNGA has largely reflected the major social, political, and economic realities and concerns of their time. It could be possible to forecast the emergence of future concerns and how much space they are going to take up in international political and social discourse.

The goal of this research was to provide an alternative framework to think about alliance formation in international politics through the UNGA resolutions that would not rely on country-level mass generalizations. Extracting issue areas from the UN resolutions by leveraging topic models, and visualizing UN voting blocs with network graphs provide us with such a framework and allow us to obtain more granular information on alliance formation in a highly dynamic and complex multidimensional policy space. The accuracy of the graph representation of interstate alliances can be validated by subject matter expertise and deeper qualitative analysis. Throughout the study, I have shown the main issue areas mainly discussed by the UNGA ambassadors and, more importantly, the noticeable differences in alliance structures within these issue areas and different time periods. This was a display of the possibility of capturing a more granular and precise international relations framework both in bilateral and multilateral diplomacy. The graph visualizations of UNGA voting data allowed us to the communities and blocs in international policy areas without sacrificing precision for simplicity. By visually and interactively inspecting such graph representations of voting patterns, a high-level understanding of multipolarity within specific issue areas such as sustainable development or nuclear disarmament can be gauged. It can be argued that network visualizations cannot provide the viewer with substantial information on the mechanisms at play in alliances and polarizations. However, it can potentially improve on the now outdated groupings of countries that fail to stand against the test of time, ignore multidimensionality by assuming all hostilities and diplomatic friendships are transferrable across policy domains, and often grouping countries together based solely on shared borders or economic maturity.

The code to collect, clean, and pre-process the data, and produce the outputs presented in this paper is available at https://github.com/egemenpamukcu/unga-resolutions.

Works Cited

Beli, D., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 993-1022.

Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., & Blei, D. M. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. In *Neural Information Processing Systems.* Vancouver.

Chuang, J., Manning, C., & Heer, J. (2012). Termite: Visualization Techniques for Assessing Textual Topic Model.

Easterly, W. (2009). How the Millennium Development Goals are Unfair to Africa. *World Development*, 26-35.

Fruchterman, T. M., & Reingold. (1991). Graph drawing by force-directed placement. *Software: Pracitce and Experience*, 1129-1164.

Holmes, K. R. (1993). New World Disorder: A Critique of the United Nations . *Journal of International Affairs*, 323-340.

Huntington, S. P. (2000). The Clash of Civilizations. In S. P. Huntington, *The Clash of Civilizations?* New York, New York: Palgrave Macmillan.

Kacowicz, A. M. (2007). Globalization, Poverty, and the North-South Divide. *International Studies Review*, 565-580.

Kloß, S. T. (2017). The Global South as Subversive Practice: Challenges and Potentials of a Heuristic Concept. *The Global South*, 1-17.

Lees, N. (2020). The Brandt Line after forty years: The more North–South relations change, the more they stay the same? *Review of International Studies*, 85-106.

Macon, K. T., Mucha, P. J., & Porter, M. A. (2012). Community structure in the United Nations General Assembly . *Physica A: Statistical Mechanics and its Applications*, 343-361.

Porter, M. A., Mucha, P. J., Newman, M. E., & Warmbrand, C. M. (2005). A network analysis of committees in the U.S. House of Representatives. *The Proceedings of the National Academy of Sciences (PNAS)*.

Sakwa, R. (2019). The International System and Models of Global Order. *Russia in Global Affairs*, 8-24.

Sievert, C., & Shirley, K. E. (2014). LDAvis: A method for visualizing and interpreting topics. *Workshop on Interactive Language Learning, Visualization, and Interfaces at the Association for Computational Linguistics.*

Smith, R. W. (2007). An Overview of the Tesseract OCR Engine. *Proceedings of the Ninth International Conference on Document Analysis and Recognition* (pp. 629-633). IEEE Computer Society.

The World Bank. (2010, April 14). Old concept of “Third World” outdated, Zoellick says. Washington, DC, United States of America.

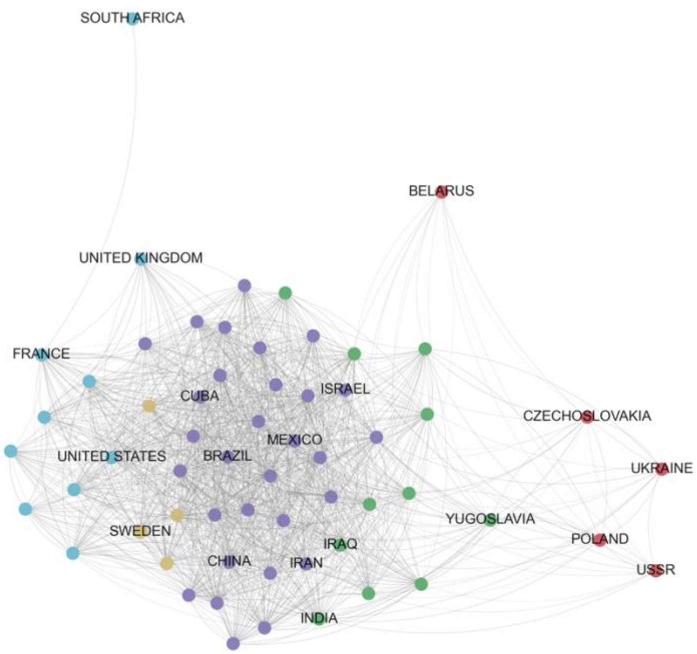
Traag, V. A., & Bruggeman, J. (2009). Community detection in networks with positive and negative links . *Physical Review E*.

van der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 2579-2605.

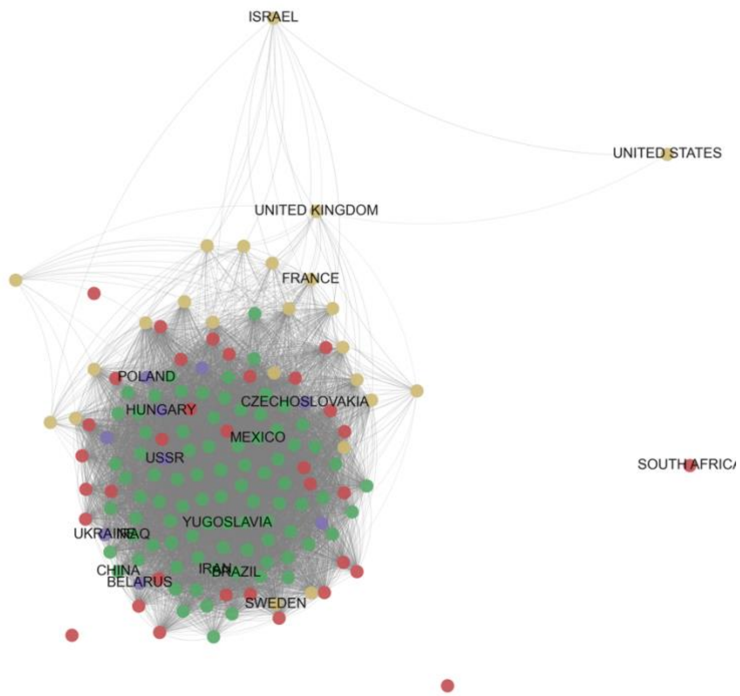
Voeten, E. (2000). Clashes in the Assembly . *International Organization*, 185-215.

Waugh, A. S., Pei, L., Fowler, J. H., Mucha, P. J., & Porter, M. A. (2011). Party Polarization in Congress: A Network Science Approach . *Physics and Society*.

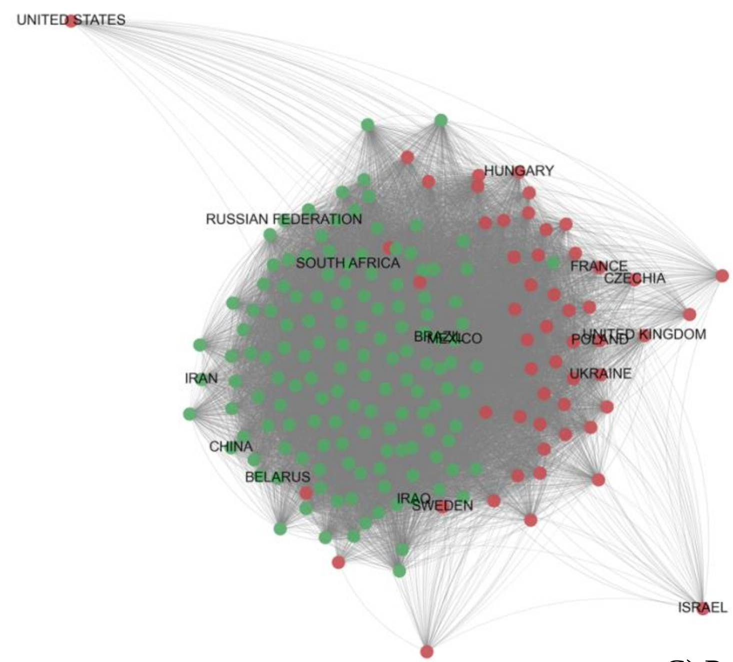
Appendix 1: Network visualizations for the Discrimination and Violence Against Women and Children issue area



A) Early Cold War Period

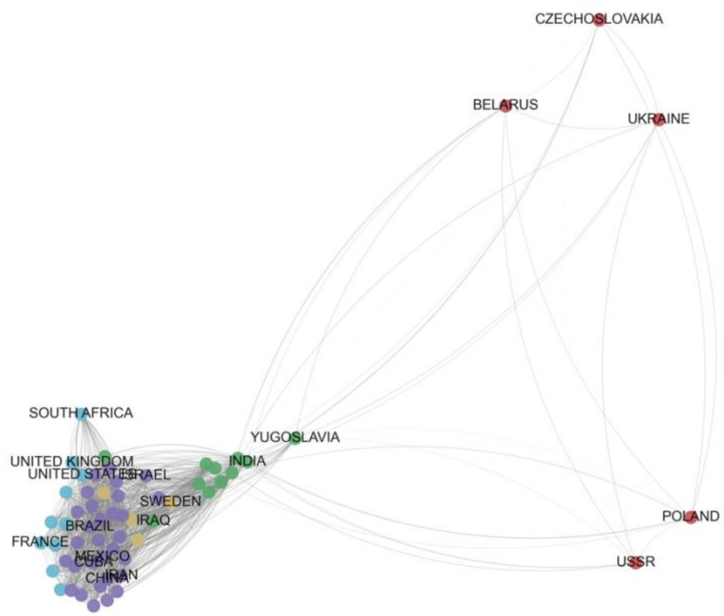


B) Transitional Period

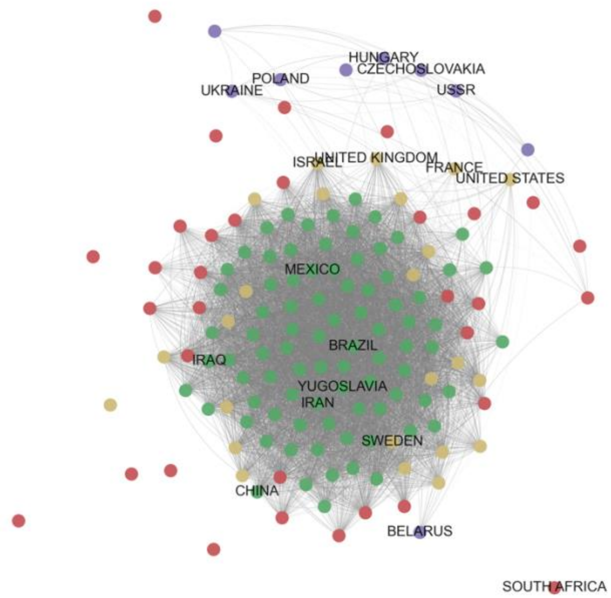


C) Post-Cold War Period

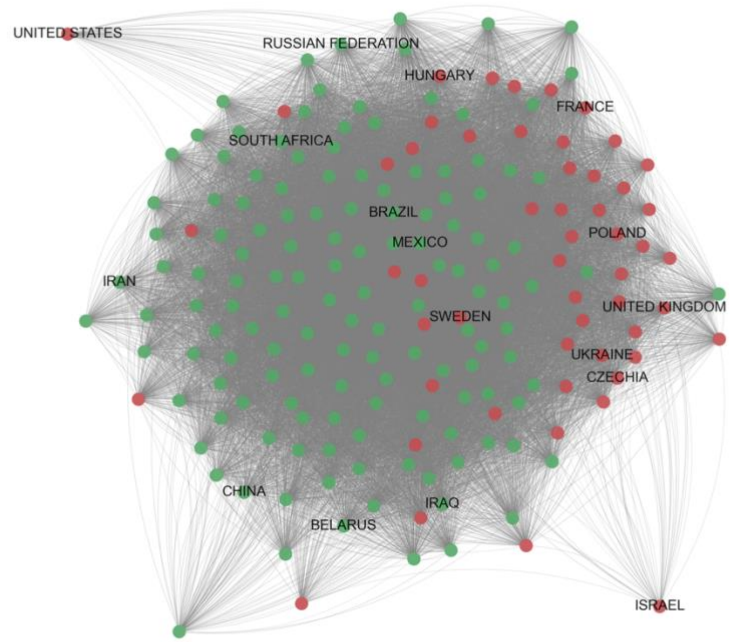
Appendix 2: Network visualizations for the UN Appointments and Administration issue area



A) Early Cold War Period

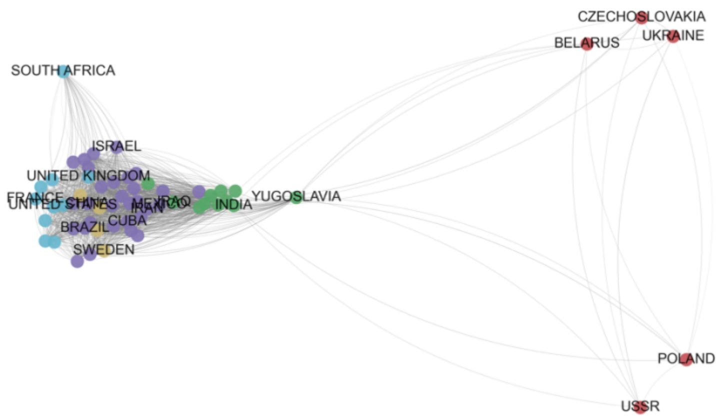


B) Transitional Period

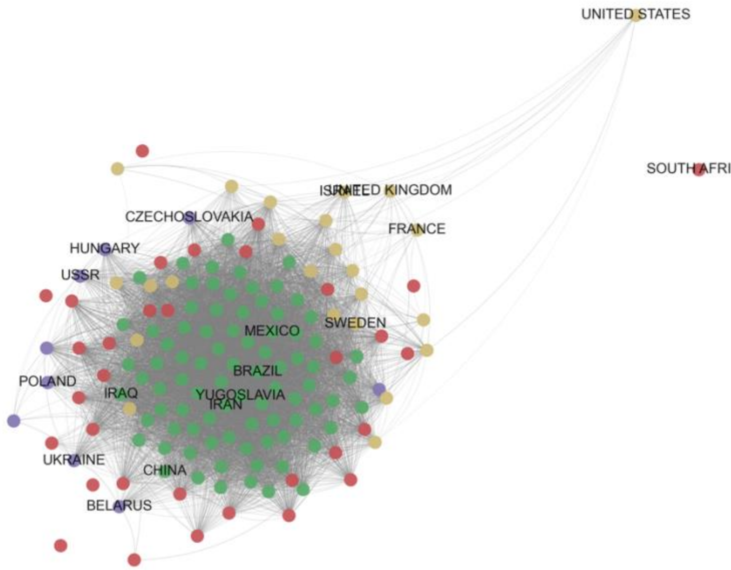


C) Post-Cold War Period

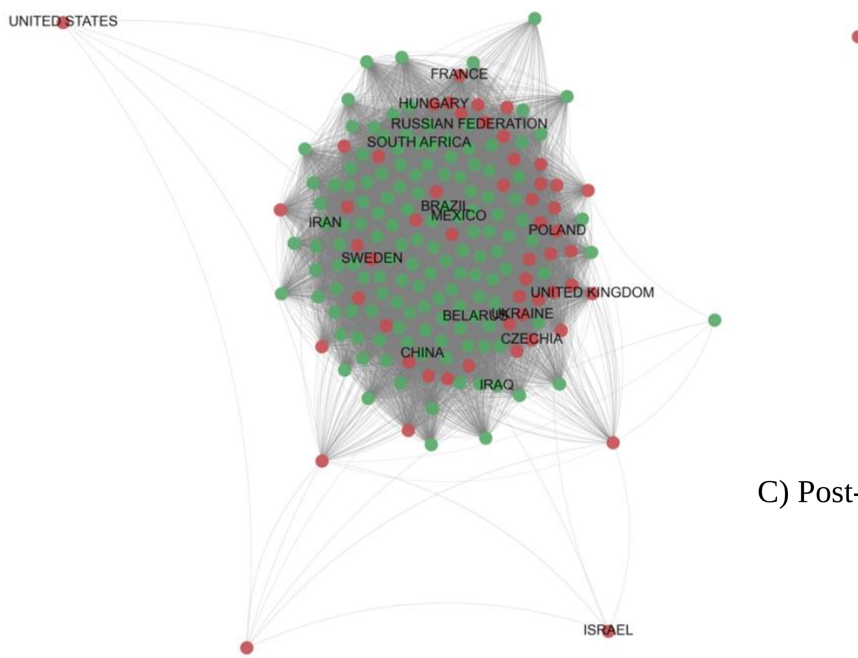
Appendix 3: Network visualizations for the UN Memberships issue area



A) Early Cold War Period

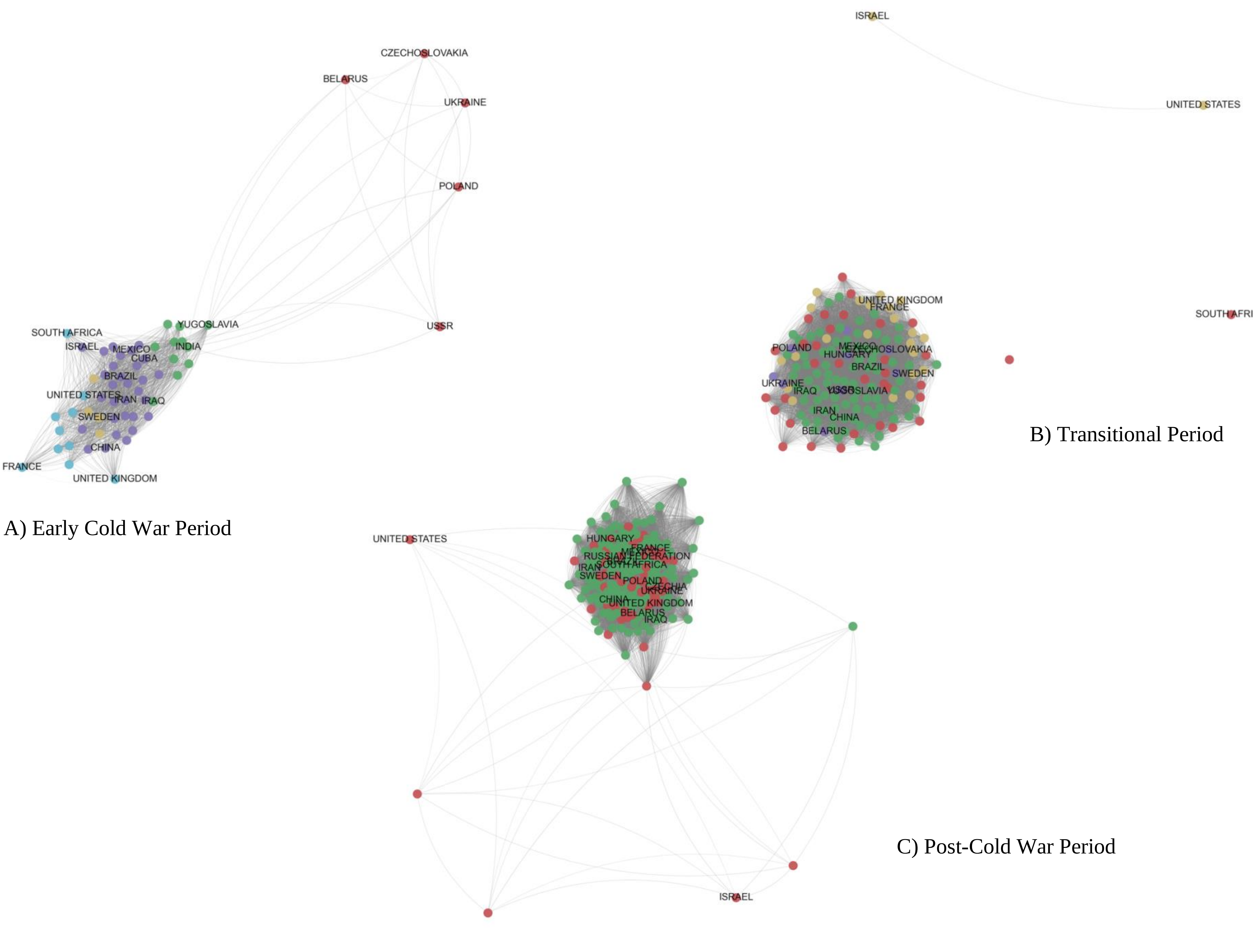


B) Transitional Period

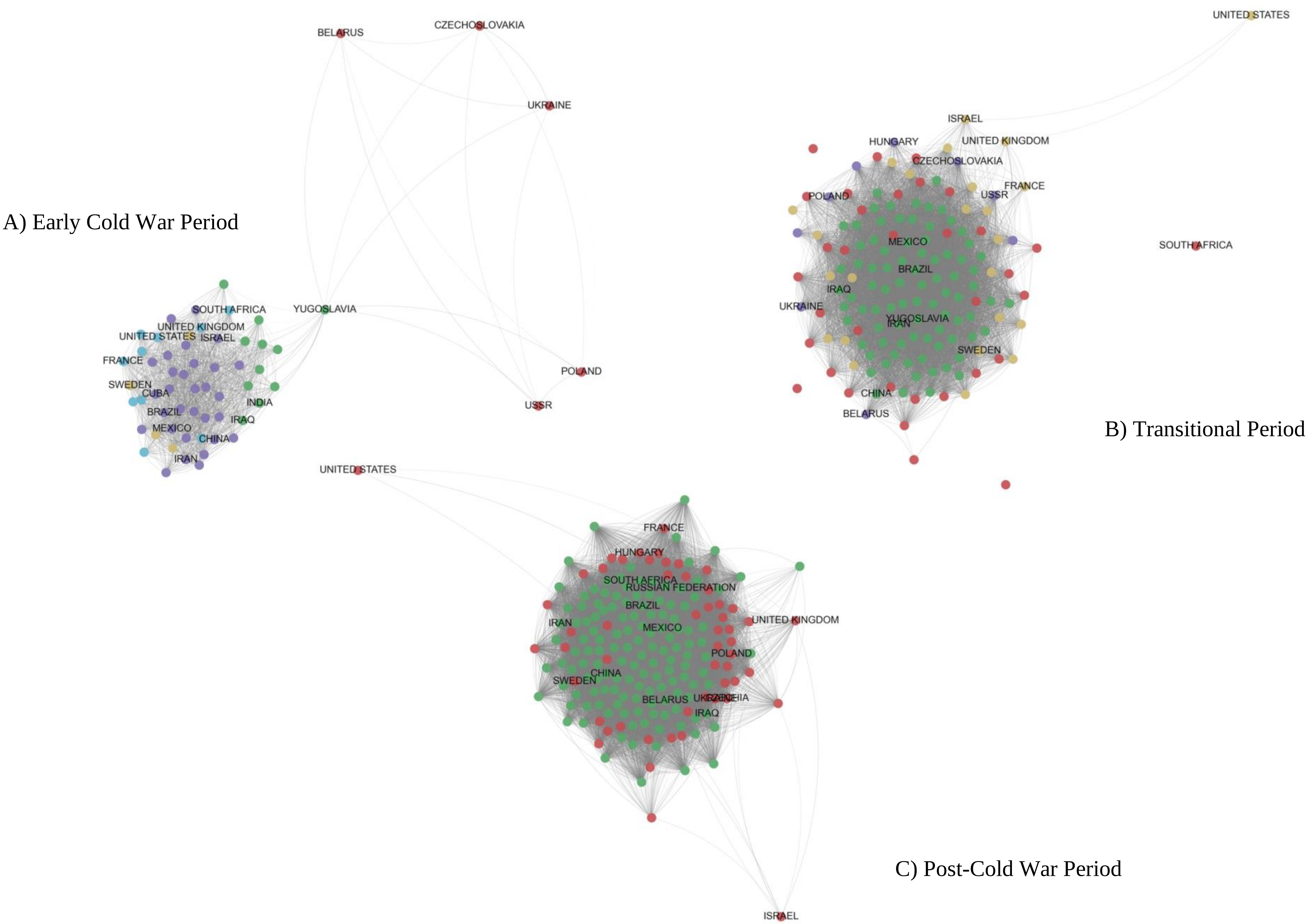


C) Post-Cold War Period

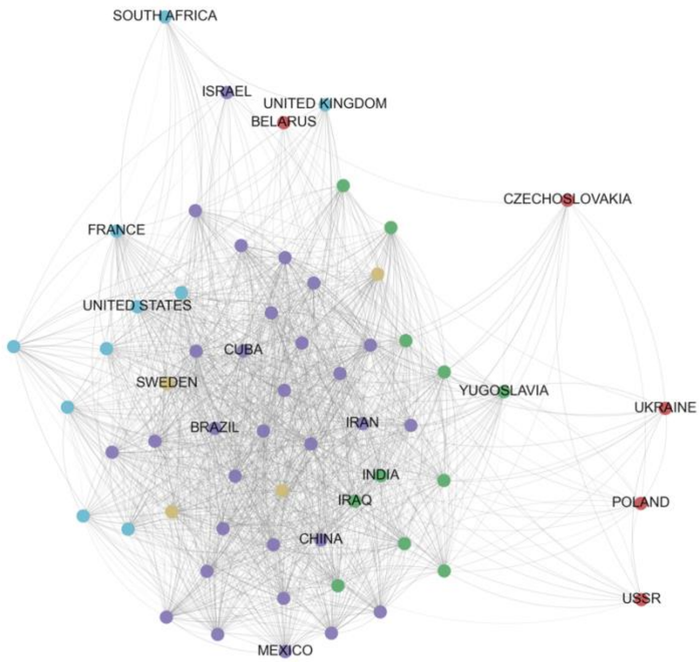
Appendix 4: Network visualizations for the Arab-Israeli Conflict issue area



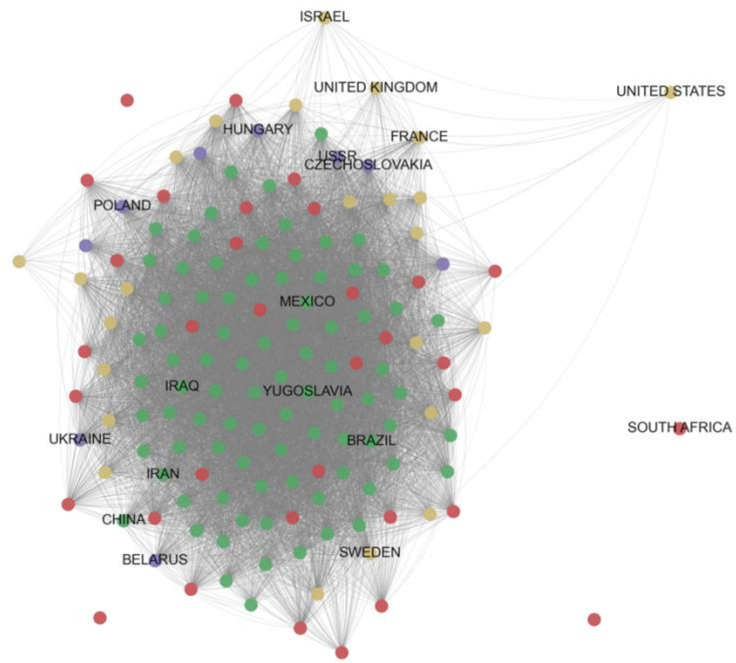
Appendix 5: Network visualizations for the Outer Space, Science and. Technology issue area



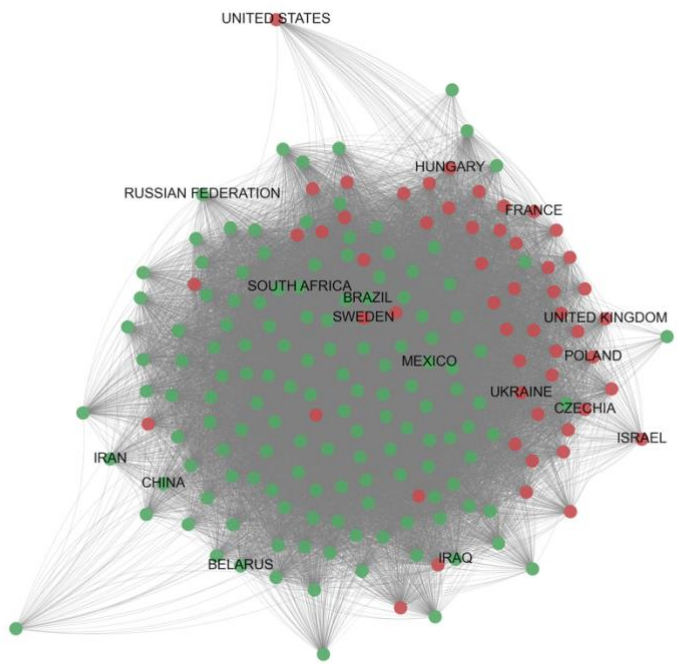
Appendix 6: Network visualizations for the Fight Against Global Crime issue area



A) Early Cold War Period

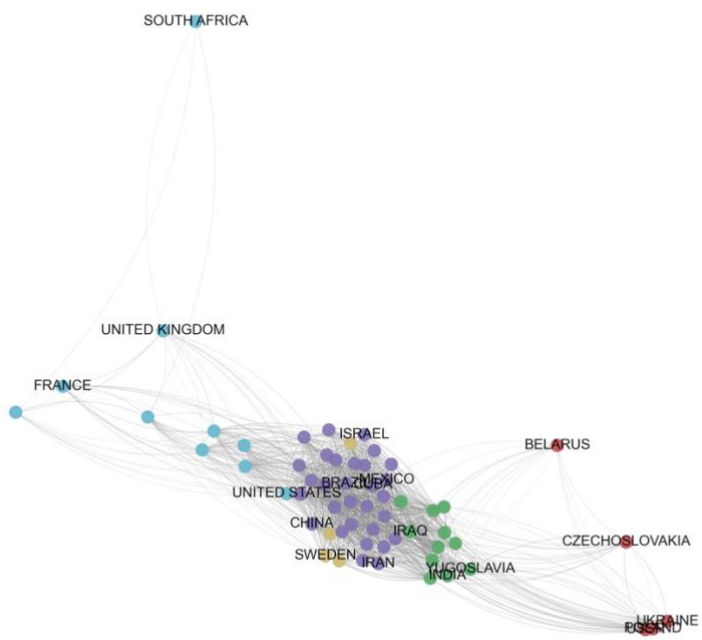


B) Transitional Period

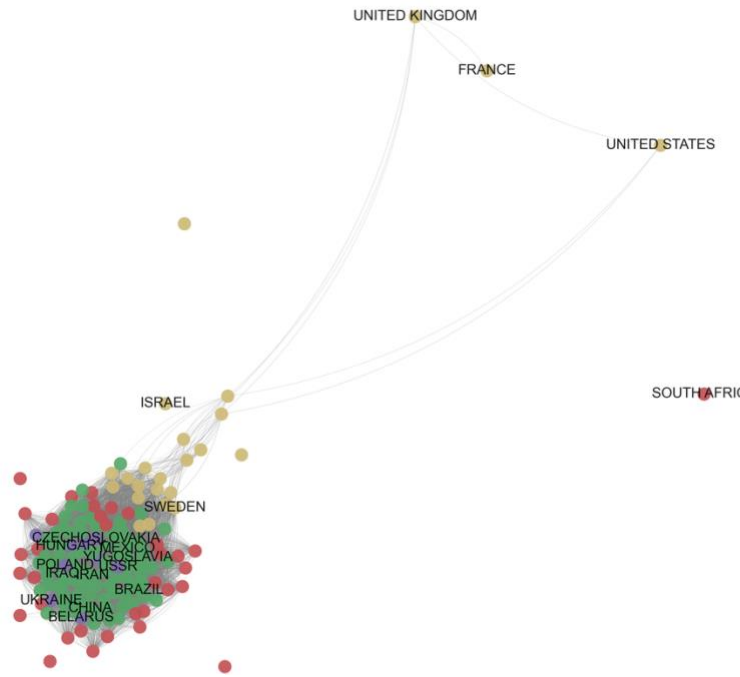


C) Post-Cold War Period

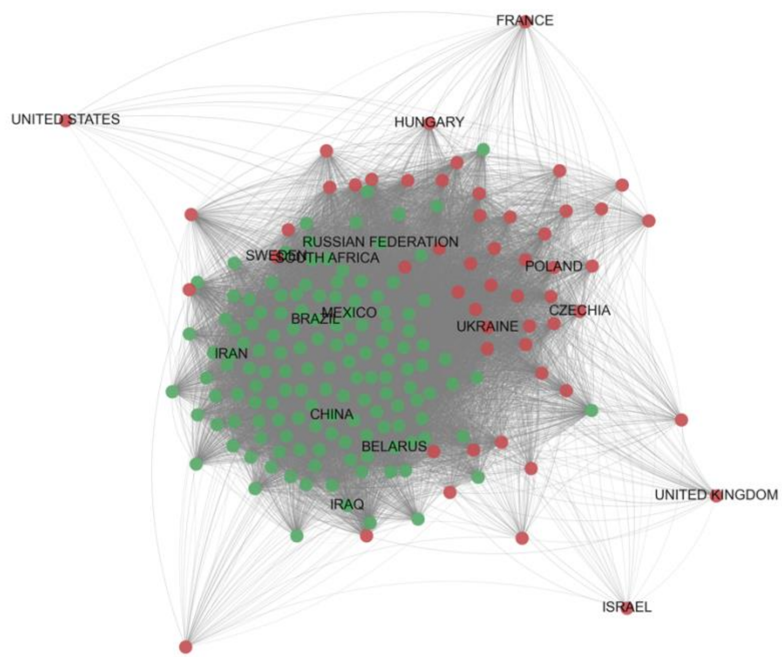
Appendix 7: Network visualizations for the South Africa and Apartheid issue area



A) Early Cold War Period

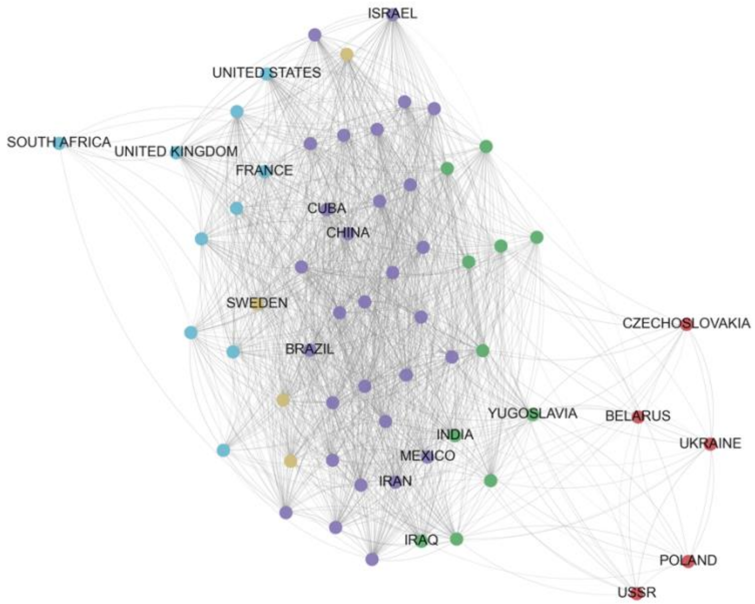


B) Transitional Period

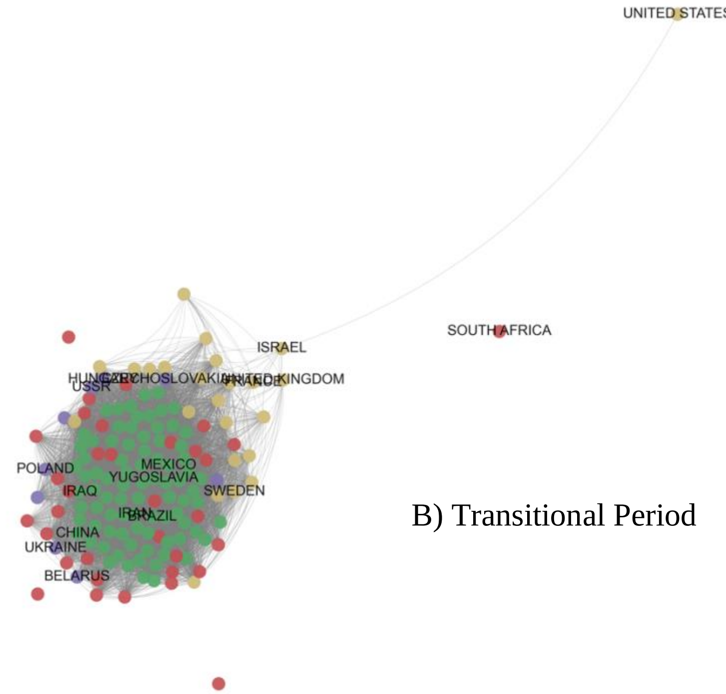


C) Post-Cold War Period

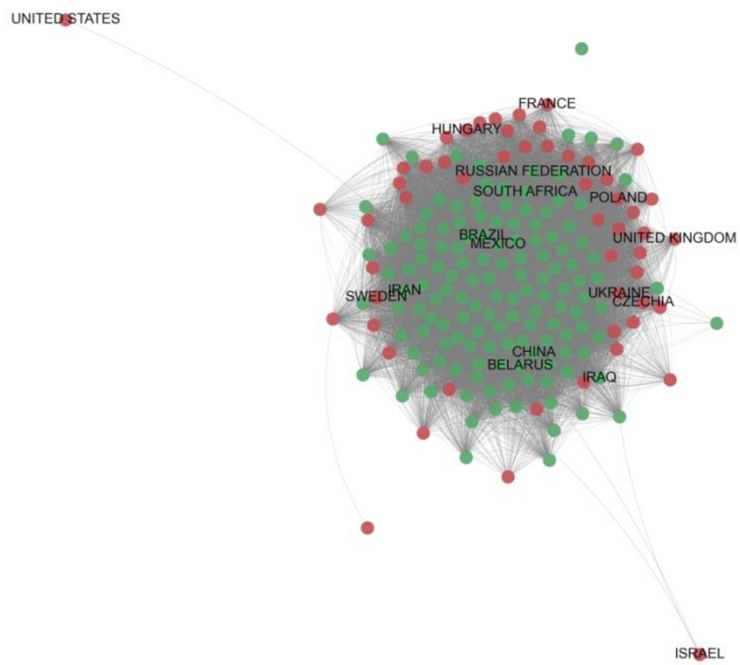
Appendix 8: Network visualizations for the Industrial Development issue area



A) Early Cold War Period

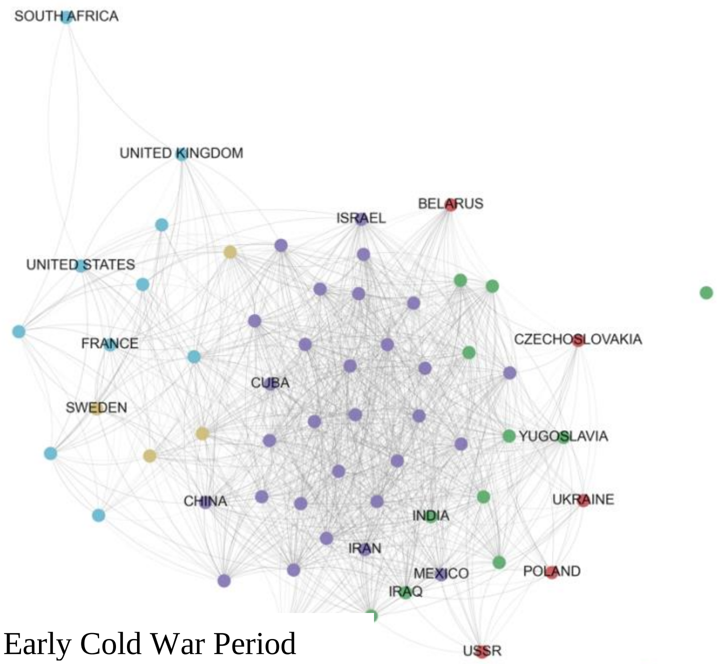


B) Transitional Period

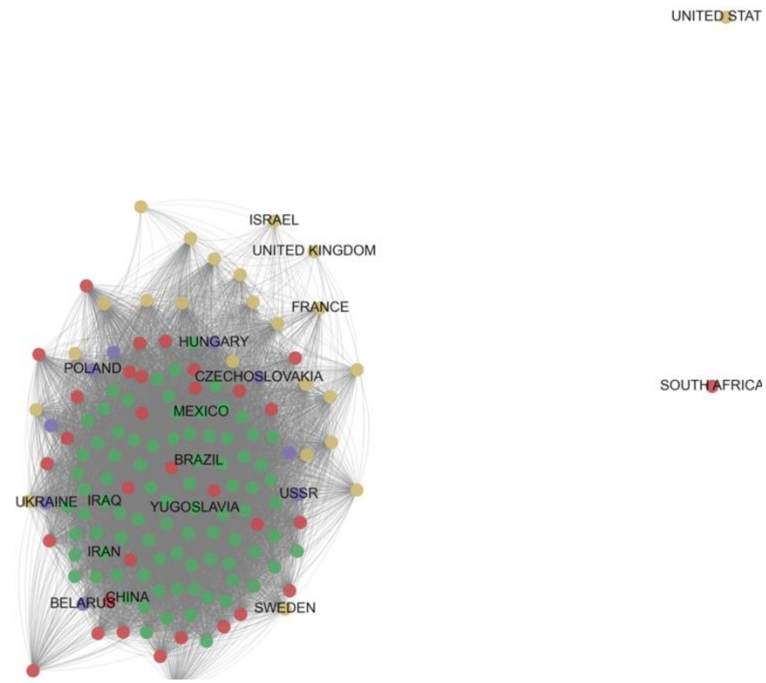


C) Post-Cold War Period

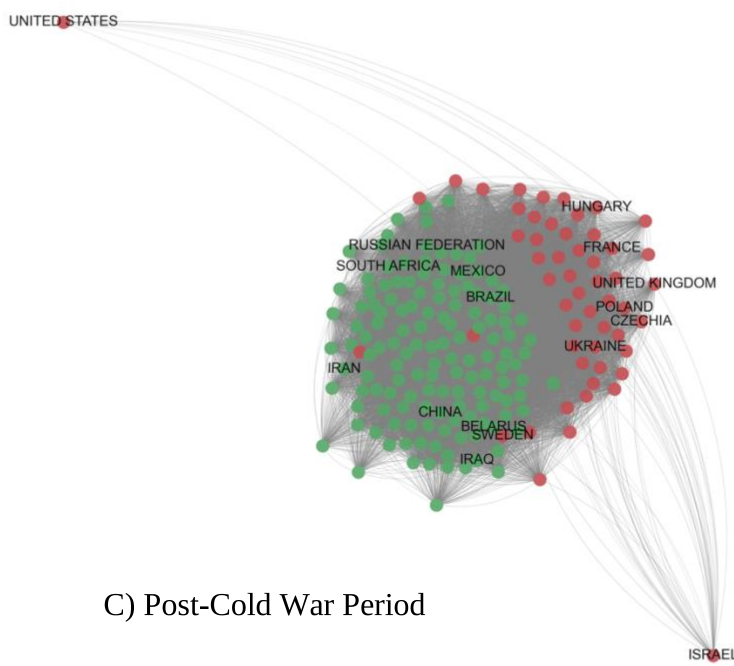
Appendix 9: Network visualizations for the Sustainable Development issue area



A) Early Cold War Period

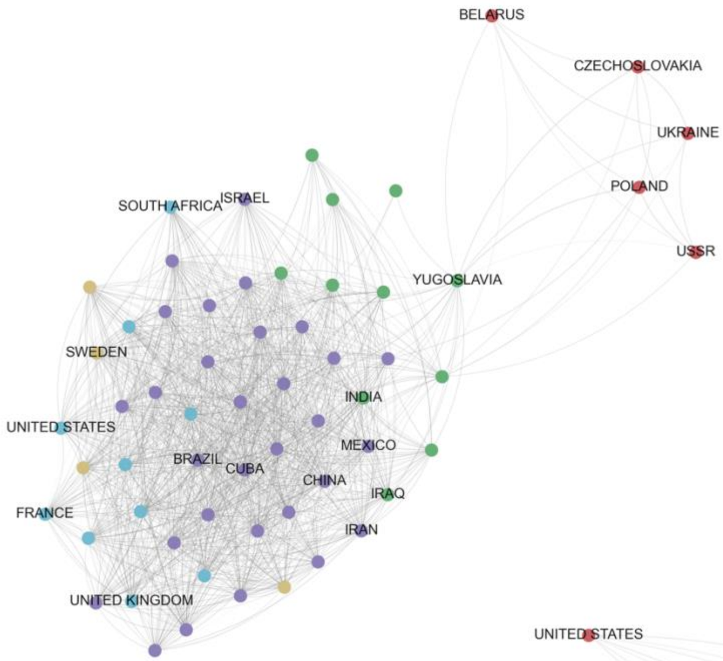


B) Transitional Period

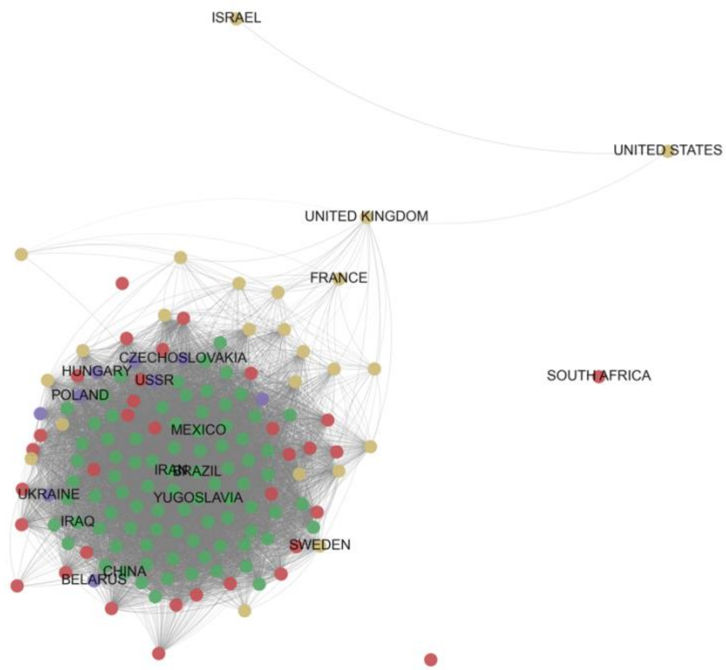


C) Post-Cold War Period

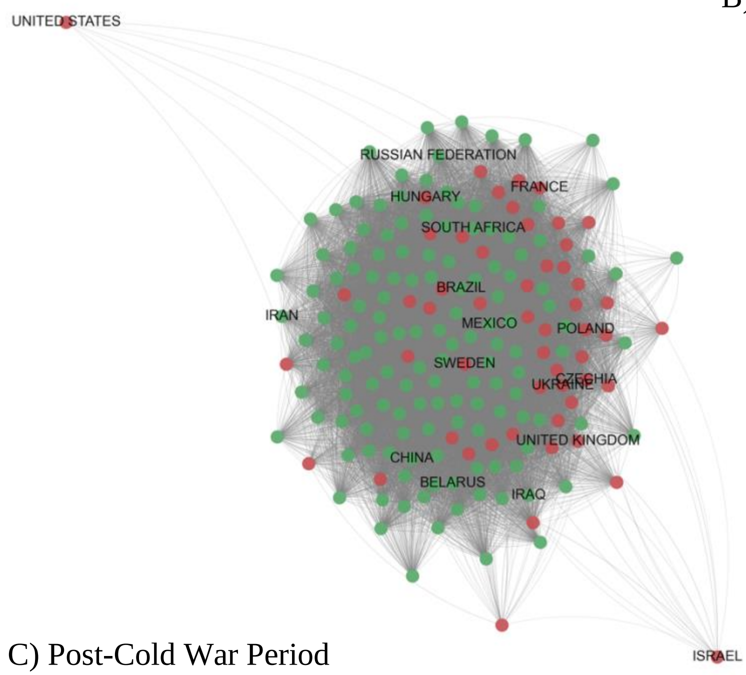
Appendix 10: Network visualizations for the Oceans and the Law of the Sea issue area



A) Early Cold War Period



B) Transitional Period



C) Post-Cold War Period