# Forecasting Defense Alliances Using Graph Theory

## Outline

I. Introduction (1.5 pages)

Research Question: How can graph theoretic models enable better strategic insights into the evolving nature of defense alliances between countries?

II. Methodology (3 pages)

III. Literature Review (4 pages)

IV. Research (27 pages)

a. Model Overview
b. Negative Ratios
c. Alliance Trends
d. Overall Model Results
e. Model Accuracies
f. Model Inaccuracies
g. Notable Dips in Accuracy
h. Future Features
    i. Domestic Features
    ii. Economic Features
    iii. International Features
    iv. Military Features
i. Sensitivity Tests

## I: Introduction

Graph theory, a branch of mathematics, holds great potential for providing strategic insights into social science and national security problem sets. The following research will explore how graph theoretic models can enable better strategic insights into the evolving nature of defense alliances between countries. Historical data reveals various patterns that characterize a shifting landscape of interstate defense alliances, with major shifts often occurring after periods of strife or change such as world wars or the break-up of countries. Expressing this historical data as a temporal graph (i.e. countries as nodes and extant defense alliances as edges) and using graph models to predict the formation or dissolution of such partnerships would enable strategic insights into how various world events and country characteristics shape the dynamic nature of defense alliances.

This research will focus on two primary strategic insights. First, the model will identify what features (e.g. proximity, regime type, religion, history of wars) are most significant for shifts in defense alliances. Second, the project's framework will demonstrate that graph theory can predict the formation and dissolution of defense alliances (i.e. the additions/deletions of graph edges). Furthermore, this research is designed to assess the feasibility of applying this branch of mathematics to security and social science problem sets by learning patterns from data expressed as graphs.

Previous academic applications of graph theory to real-world examinations of social science problems have been conducted in a limited and often tangential manner. These have only used graph theory as a theoretical framework to model alliance relationships between countries but have not used graph theory – along with data and machine learning – to predict these alliances. This project will help bridge this gap. Additionally, this research will help pioneer the assimilation of technologies from other spheres of research into national security, strengthening policymaking through data-driven approaches and evidence-based decision-making. Furthermore, in addition to illustrating interstate alliances, graph-based entity relationships are typical in fields such as terrorism networks, illicit trafficking networks, organized crime syndicates, and intelligence networks. The success of this research will enable the application of this unusual partnership of two historically disparate fields (mathematics and social science) to problem sets vital for national security.

## II: Methodology

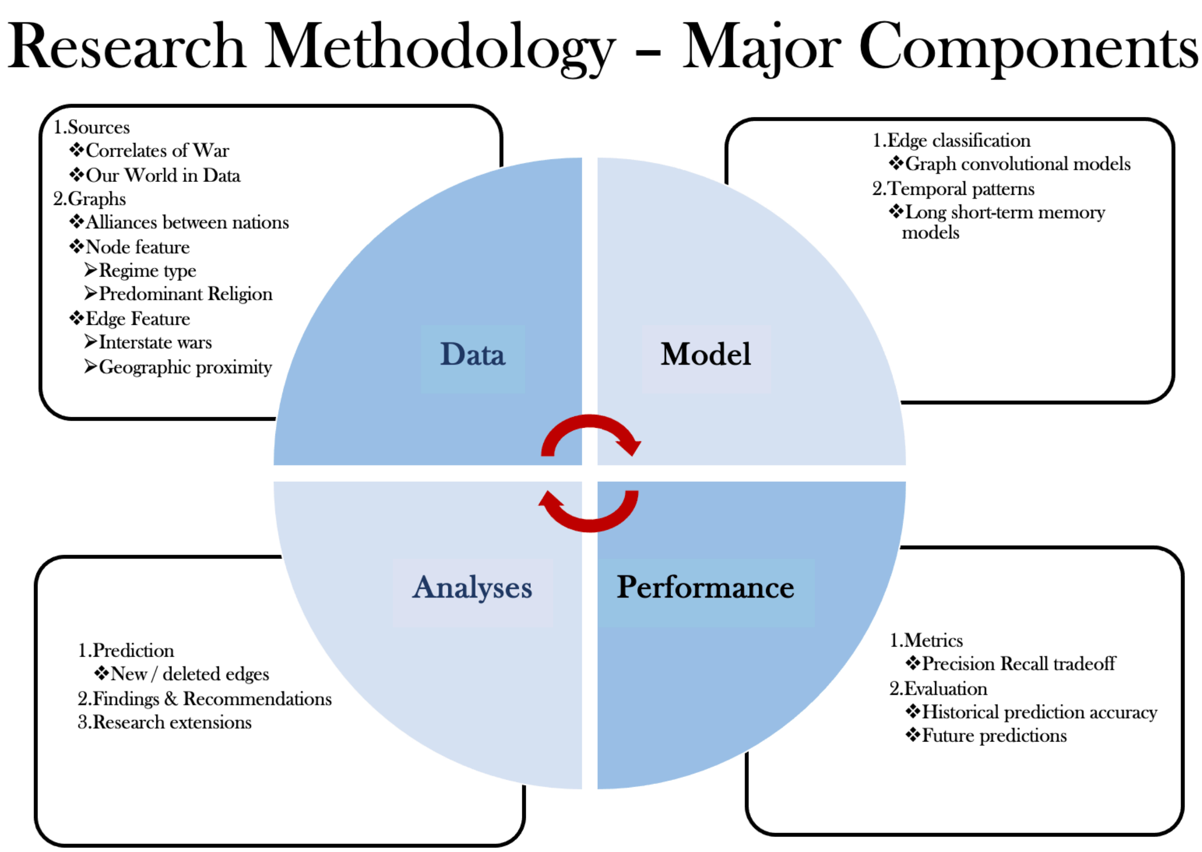


*Figure 1: Methodology Summary*

The research uses publicly available online data sets detailing defense alliance data, war data, regime data, geographic proximity data, and religion data from 1816 to 2012. The model relies on the Correlates of War project, which is peer-reviewed and curated by social scientists, and the Our World in Data project, which is produced by Oxford University researchers and the non-profit research organization Global Change Data Lab.[1][2][3][4][5]

The study consists of an edge prediction model that predicts the likelihood of an edge (i.e. a defense alliance) forming or disappearing between two nodes (i.e. countries) using temporal, historical data. The model leverages this data to forecast defense alliances through a graph convolutional network (GCN) that integrates a long short-term memory (LSTM) network.

First, the data is divided into yearly segments that are used to construct a sequence of yearly static graphs with nodes and edges. Each node represents a country, and each edge depicts a recorded defense alliance between countries. Each country node will have features describing country characteristics (e.g., regime, predominant religion). Each edge will have features describing interstate relationships (e.g., wars, geographic proximity). Using this node and edge data, a static graphic is created to depict the alliances of each year. The data sets provide information for 196 years. A temporal progression of graphs is modeled by considering these static graphs in sequence. A successful model will predict the static graph for each year using only the previous years' static graph data (e.g., predicting what 2000 will look like by using the 183 static graphs that depict 1816 to 1999).

After constructing the graphs, the model is trained using multiple layers. The model employs supervised learning that uses an equal number of positive edges (i.e. confirmed alliances) and negative edges (i.e. confirmed lack of alliances) for effective training. The first layer of the model is the GCN layer, which updates each node's characterizations based on its neighbor nodes' features. GCNs are particularly advantageous for capturing structural information and spatial dependencies.[6] The second layer is the LSTM layer, which processes the temporal data and captures the long-term, chronological patterns and dependencies, enabling the model's learning.[7] Sustaining this historical or temporal context through these networks increases the accuracy of edge predictions. The model learns by looking at the patterns of historical static graphs and trying to predict the current year's alliances and the absence of alliances between all pairs of countries. It repeatedly iterates through the data and refines its predictions. The last layer of the model is the fully connected layer, which yields the prediction scores for potential edges on a scale of 0.00 to 1.00, with 1.00 being 100% likelihood of forming a new alliance.

The last step is evaluation. Since there is data from 1816 to 2012, the modeler can cross-check the model's predicted results with the real-life data to evaluate how successful the model is in predicting the formation or dissolution of defense alliances. The modeler analyzes the model's results and sorts them into four bins: true positives, false positives, false negatives, and true negatives. This helps the modeler calculate its accuracy, precision, and recall levels, as well as its F1 score (harmonic mean) and AUC-PR (area under the precision/recall curve). The project's goals and problem set will dictate whether the modeler should prioritize accuracy vs. precision levels. For example, in the medical field, the risk of a false negative (i.e. lethality) is much greater than that of a false positive (i.e. more testing), so the modeler may want to prioritize recall. In other cases, the cost of a false positive may be high, requiring greater focus on higher precision. Other problem sets are not as easy to discern and often benefit from following the AUC-PR curve when deciding how to balance accuracy and precision.

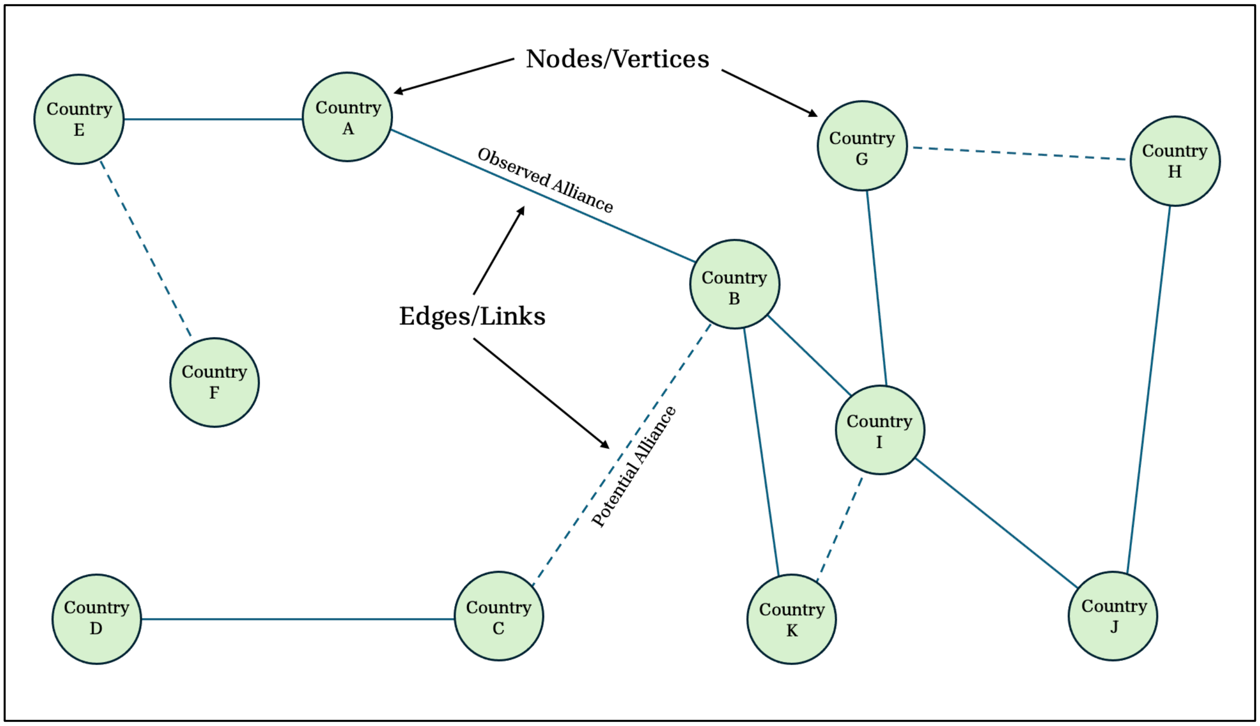


*Figure 2: Node and Edge Illustration*

## III: Literature Review

Much of the prior work focusing on graph theory emphasizes social science fields as potential applications for the theory. However, many of these studies do not illustrate this theory through practical case studies. Furthermore, the few studies illustrating a case study do not create a predictive tool for dynamic networks where relationships change and evolve. This research will apply temporal graph edge classifications to the study of defense alliances to help bridge this gap.

First, this section will discuss theoretical works that exclude explicit case studies. Then, it will examine graph theory approaches that are too simple for practical application to security problem sets. Next, it will examine theoretical works that contain recommendations on how to improve application. Finally, it will analyze an application study that omits a predictive mechanism.

Several authors provide overarching insights into network science and graph theory. Frank Harary's 1969 book provides one of the original mathematical basics of graph theory and matrices.[8] Harary and Norman's follow-on research states that graph theory should be used to depict interpersonal relationships in social science fields.[9] However, it omits explicit social science case studies and largely serves as an introductory text into graph theory for future social scientists to refer to when conceptualizing their own projects.

Other scholars' works have also remained theoretical and excluded explicit case studies. Balakrishnan and Ranganathan's 2012 book discusses the basic principles of graph theory and examples of how graph theory can be applied to the hard sciences, psychology, sociology, and computer science.[10] Brigham, Dutton, Haynes, and Hedetniemi's research uses proofs, corollaries, and theorems to detail the mathematical properties of alliances.[11] Similarly, Lal's work discusses the mathematical properties of alliances and specifies that alliances with minimum cardinality are the most notable.[12]

The field took a leap forward when Harary's 1953 work was the first to present the concepts of a signed graph and balance.[13] Signed social networks label each edge as a positive/ally or negative/adversarial relationship. A signed graph is balanced if the product of the edge signs surrounding each cycle in the graph is positive. Harary and Kabell expand on these ideas with their 1980 paper positing a simple algorithm called a linear time algorithm that can be used to test any signed graph for balance.[14] Chiang, Hsieh, Natarajan, Dhillon, and Tewari further support the use of signed networks for social network analysis.[15] Doreian and Mrvar discuss how studies examining signed social networks have overlooked universally positive nodes and nodes with positive links to pairs of nodes that have negative links or are in adversarial subsets.[16] However, while signed graphs can be helpful for many other social science problem sets, there is often no standardized way to classify nation-state relationships as positive or negative. Many state-actor relationships are too nuanced to categorize as strictly allied or strictly adversarial in every sector. States allied on defense may disagree on climate agreements, and adversarial states may team up against counternarcotics. Consequently, the proposed model uses unsigned edges.

Kumar, Spezzano, Subrahmanian, and Faloutsos's 2016 piece expanded on previous literature by adding a "weight" element to signed relationships and a "goodness" element and "fairness" element to node characterization.[17] "Goodness" depicts to what extent the node is trusted or liked by other nodes. "Fairness" depicts to what extent the node is fair in rating other nodes' "goodness." The proposed model also uses node features to improve edge prediction.

As study in the field progressed, several authors began to examine machine learning in graph theory. Ahmad, Akhtar, Noor, and Shahnaz's 2020 piece proposes a new link prediction algorithm based on centrality and common neighbor nodes.[18] The paper illustrates how the count of common neighbors – or neighboring nodes and links – influences additional links. If this approach were applied to the proposed model depicting defense alliances, the algorithm would use neighboring links for other countries as indications for additional linkages. For example, consider a scenario in which the model attempts to discern whether the United Kingdom and France are linked in an alliance in a given year. If the United Kingdom has an alliance with the United States and France has an alliance with the United States, then that common link to the United States may be indicative of a relationship between the United Kingdom and France. This is largely a database approach, leading to a cruder approximation. Practically, perhaps one country needs several links, and perhaps another country needs only a few links. Rather than only counting the edges, the proposed model will learn the history and complex relationships between different types of countries through node and edge features, using the same data in a more sophisticated way.

Other scholars' works have expanded on previous literature by identifying strategies to improve graph theory's application to social science fields. Lichtenwalter and Chawla provide solutions to link prediction's inherent evaluation limitations, recommending the use of random edge sampling, varied thresholds, and precision-recall threshold curves and associated areas in place of receiver operating curves.[19] Lichtenwalter and Chawla partnered with Yang on a paper expanding on precision-recall threshold curves and the top K predictive rate, as well as illustrating what a strong precision-recall curve looks like.[20] Che, Wang, and Xu detail how social networks are dynamic and change over time, which makes predicting the addition or deletion of links challenging.[21] They recommend using deep learning, specifically convolutional neural networks and long short-term memory networks. Shafique proposed a data clustering algorithm to identify existing but undiscovered alliances in graphs.[22] Doreian discusses how modelers can use blockmodeling: a process that simplifies social networks by clustering similar nodes into blocks.[23] Doreian also advocates for the use of multiple indicators, rather than one indicator, to depict the nuances of social networks. This multiple-indicator approach is particularly helpful for the proposed model, as it uses multiple edge and node features to improve its predictive success. Furthermore, Cranmer, Desmarais, and Kirkland posit a theory that, in addition to characteristics largely accepted in the literature, changes in alliance graphs are primarily dependent on the graphs' structures.[24]

Doreian and Mrvar's 2015 paper provides a rare case study applying graph theory, particularly blockmodeling, to illustrate a signed graph.[25] Their graphs identify counts, communities, and clusters of countries using Correlates of War data sets. However, it does not create a predictive tool.

Overall, the literature essentially contains theoretical works, graph theory approaches that are too simple for practical application to security problem sets, and theoretical works with recommendations on how to improve application. These studies do not illustrate application through a predictive case study with dynamic networks where relationships evolve. The proposed model will apply temporal graph edge classifications to defense alliances to help bridge this gap. The next step is applying the above concepts (random edge sampling, varied thresholds, precision-recall curves, top K predictive rate, convolutional neural networks, long short-term memory networks, clustering and blockmodeling, and multiple indicators) to a temporal, predictive model that can predict alliances using social network analysis and data from the Correlates of War project and the Our World in Data project.[26][27][28][29][30]

## IV: Research

### a. Model Overview

The model leverages graph theory, machine learning, and online datasets to forecast the likelihood of edges on a static graph of a target year using the static graphs from previous years. Examples of yearly static graphs are depicted below. As illustrated, the number of nodes (i.e. states) and edges (i.e. defense alliances) increases as the world evolves from earlier to later years, modern years.

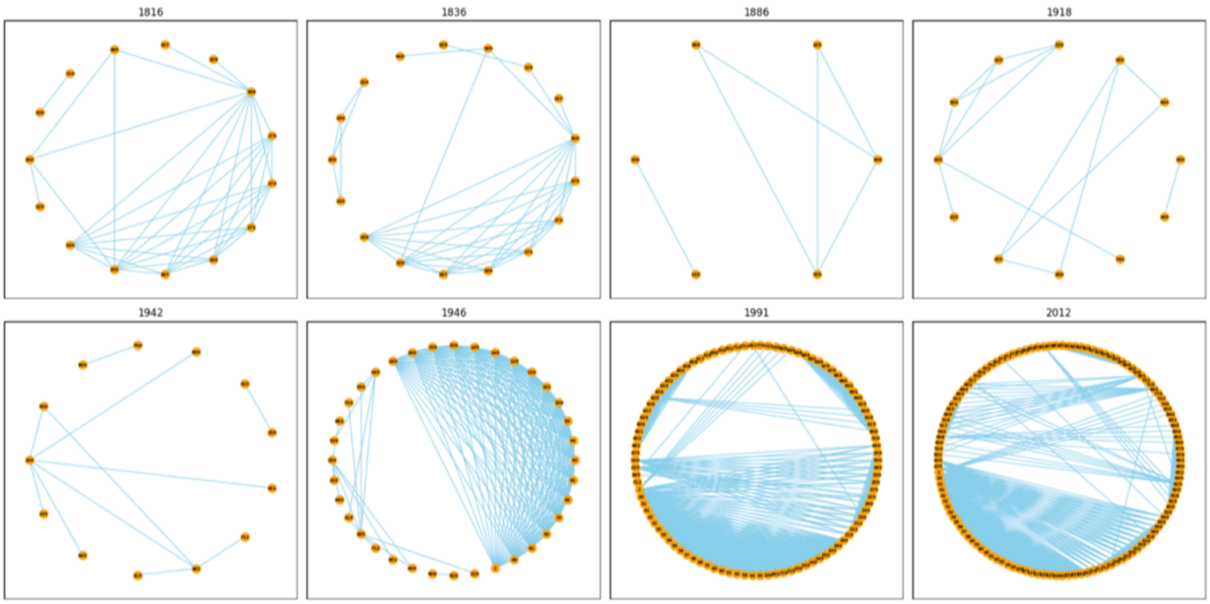


*Figure 3: Static Graphs for 1816, 1836, 1886, 1918, 1942, 1946, 1991, and 2012*

Node and edge features are obtained using online data sets. The initial model used two data sets – regime type and religion – for the node features. The model also used two data sets – proximity and interstate wars – for the edge features and prioritized characteristics such as edge weights and connection frequency. Finally, the model used an interstate defense alliance data set to provide the static graph data and validate the model's results. The model's objective is to capture temporal patterns and structural dynamics using these historical data sets, a Graph Neural Network machine learning framework, and supervised learning (as elaborated on in the literature review).

### b. Negative Ratios

Negative ratios play a vital role in training algorithms for supervised learning tasks like edge prediction. The negative ratio establishes how many non-existent edges (i.e. the lack of a defense alliance between two countries) are selected for training in relation to existing edges (the presence of a defense alliance between two countries). For the purposes of this paper, "negative samples" will refer to non-existent edges, and "positive samples" will refer to existing edges.

In graph-based predictive models, a dataset can often have far more negative samples than positive ones due to the natural sparsity of graphs. This is normal because most possible connections between nodes do not exist. However, if the number of negative samples is excessively large compared to positive samples, it can trigger overrepresentation. Overrepresentation makes it harder for the model to focus on learning the patterns that define positive samples and might skew the model's predictions toward negative outcomes. For example, in the case of this paper's model, overrepresentation might cause the model to find it difficult to predict the presence of interstate defense alliances (i.e. positive outcomes) and instead skew towards predicting the absence of interstate defense alliances (i.e. negative outcomes). It is difficult to create a model that is equally good at predicting both these positive and negative outcomes. This is why modelers typically choose whether they want to identify all the possible positive outcomes with some inaccuracies or whether they want to prioritize the accuracy of what they do identify and risk missing some positive outcomes (or somewhere in the middle). Different negative ratios are suitable for different project goals, so the modeler must also select an optimal negative ratio value best suited for the model's objective and data.

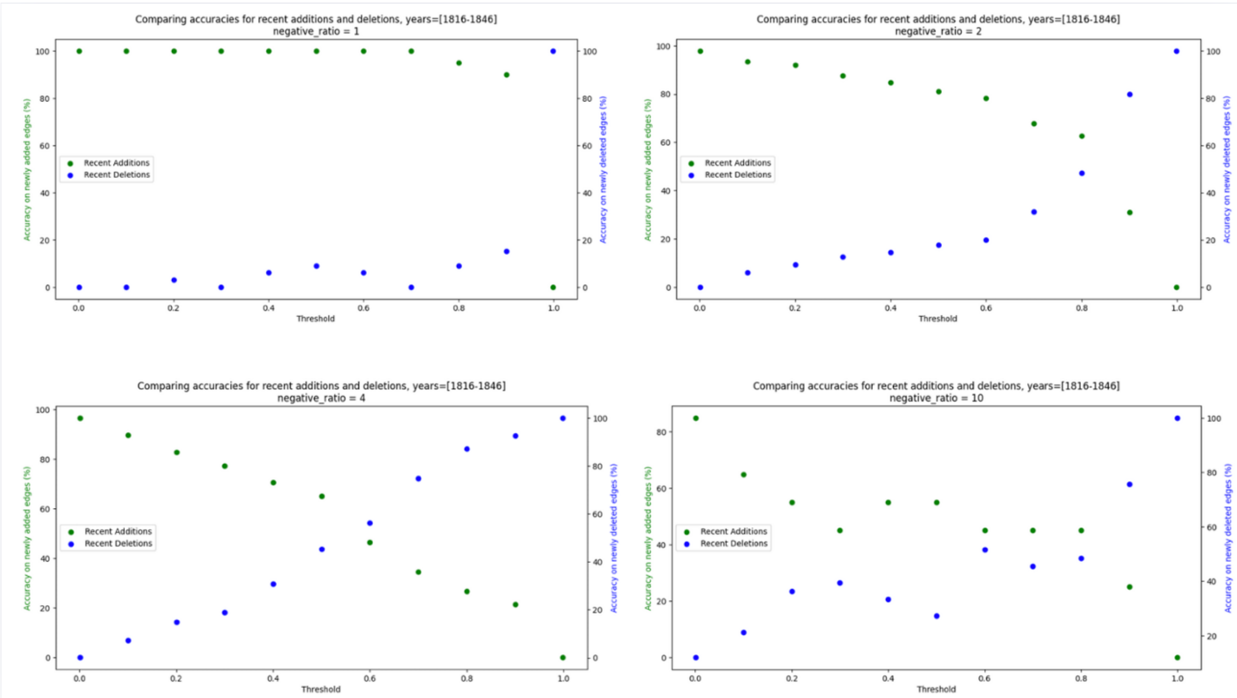


*Figure 4: Newly Added and Deleted Edge Accuracy for Negative Ratios 1, 2, 4, and 10*

A negative ratio of 1 mandates that, for every positive sample, one negative sample is selected. This type of sampling can reduce the risk of uneven training by allowing the model to learn equally from positive and negative samples. Modelers should use a negative ratio of 1 when the number of negative samples does not overwhelmingly outnumber the positive samples. The top left graph in Figure 4 compares the accuracies for defense alliance formation and dissolution predictions from 1816–1846 using a negative ratio of 1. As illustrated, a negative ratio of 1 is optimal when prioritizing the accuracy of newly added edges (depicted in green) over the accuracy of newly deleted edges (depicted in blue).

A negative ratio of 2 stipulates that, for every positive sample, two negative samples are selected. This type of sampling increases the diversity of the negative sample pool, possibly improving the model's predictive ability. Modelers should use a negative ratio of 2 when the number of negative samples are moderately plentiful. In other words, a negative ratio of 2 is optimal for graphs with moderate sparsity. The top right graph in Figure 4 compares the accuracies for defense alliance formation and dissolution predictions from 1816–1846 using a negative ratio of 2. As illustrated, a negative ratio of 2 is useful when prioritizing the accuracy of newly added edges over the accuracy of newly deleted edges for smaller thresholds, as well as prioritizing the accuracy of newly deleted edges over the accuracy of newly added edges for larger thresholds.

A negative ratio of 4 specifies that, for every positive sample, four negative samples are selected. This type of sampling helps illustrate more variability in the negative samples. Modelers should use a negative ratio of 4 when the number of negative samples are significantly plentiful. In other words, a negative ratio of 4 is optimal for graphs with considerable sparsity. The bottom left graph in Figure 4 compares the accuracies for defense alliance formation and dissolution predictions from 1816–1846 using a negative ratio of 4. As illustrated, similar to the threshold graph for a negative ratio of 2, a negative ratio of 4 is useful when prioritizing an inverse relationship between the accuracy of newly added edges over the accuracy of newly deleted edges.

A negative ratio of 10 requires that, for every positive sample, ten negative samples are selected. This type of sampling emphasizes negative samples more than positive samples, so modelers should use a negative ratio of 10 for exceptionally sparse graphs in which positive samples are rare compared to negative samples. In other words, a negative ratio of 10 would be optimal for a scenario in which the presence of interstate defense alliances is extremely rare, so the model needs to prioritize understanding the lack of interstate defense alliances. However, too high a negative ratio increases the risk of overfitting on negative samples and may skew the model's attention away from predicting positive outcomes. Consequently, selecting the optimal negative ratio is vital for preserving the model's reliability. The bottom right graph in Figure 4 compares the accuracies for defense alliance formation and dissolution predictions from 1816–1846 using a negative ratio of 10. As illustrated, model performance using a negative ratio of 10 exhibits an irregular inverse relationship between the accuracy of newly added edges over the accuracy of newly deleted edges.

Selecting the optimal negative ratio for the problem set is key for preserving the model's purpose and accuracy. There are two considerations that modelers should remember when selecting negative ratios: computational complexity and graph sparsity. First, higher negative ratios increase the size of the training dataset, requiring more computational (memory and processing) power. Second, higher negative ratios are optimal for sparser graphs to guarantee proper negative sample representation.

Comparing the results for negative ratios of 1, 2, 4, and 10 reveals two noteworthy characteristics. First, as the threshold increases, the performance on the newly deleted edges generally increases and the performance on the newly added edges generally decreases. Second, as the negative ratio increases, the performance on the newly deleted edges generally increases and the performance on the newly added edges generally decreases. In summary, selecting the optimal negative ratio is critically important for model performance analyses and interpretation.

### c. Alliance Trends

As part of understanding the historical evolution of alliances, the study first analyzed the overall alliance trends mined from the interstate defense alliance dataset. As illustrated below, countries are more likely to be a part of a defense alliance in later, modern years than earlier years. Also, the apparent sharp modern rise in alliance counts is easily explained by the concurrent rise in the number of modern countries, since the number of possible edges in a graph is proportional to the square of the number of nodes in the graph.

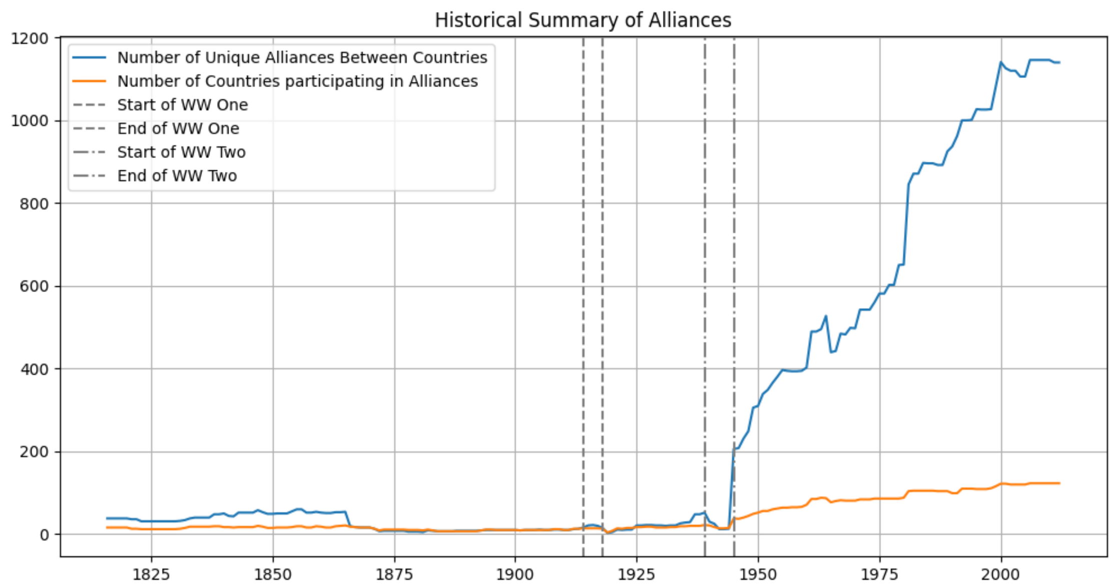


*Figure 5: Y-Axis #1: Data-Driven Visualization of Interstate Defense Alliance History*

Below is another version of the same graph with a different y-axis to discern the details better. Note that during WWI, the number of unique alliances between countries increased, and during WWII, the number of unique alliances between countries drastically decreased and only really increased towards the end of the war. This detail is important when analyzing the results of the model later in the paper.

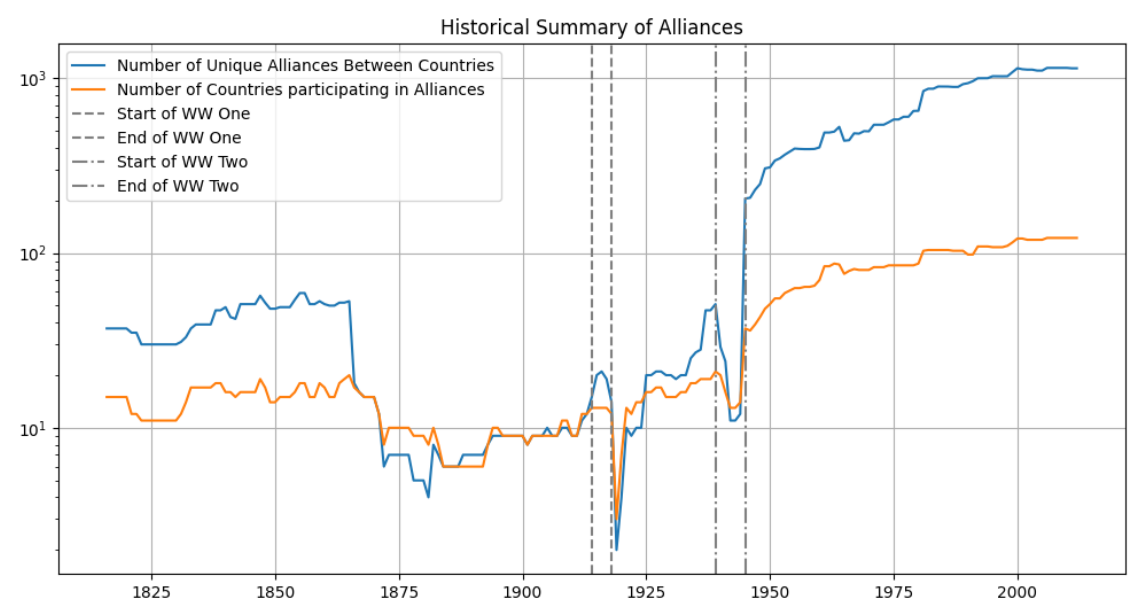


*Figure 6: Y-Axis #2: Data-Driven Visualization of Interstate Defense Alliance History*

## d. Overall Model Results

Overall, the model is accurate in terms of predicting alliances for any year given historical data, with one-off exceptions that will be examined later in this paper. Illustrated below is a graph comparing the trained model's results from 1817-2012 – depicted in blue – with a random, untrained model's results – depicted in orange. Remarkably, the model has maintained a high accuracy level as it transitions into the 21st century with globalization and the contemporary explosion of international defense agreements. The large number of modern countries and alliances does result in a small degradation in overall accuracy in the most recent years.

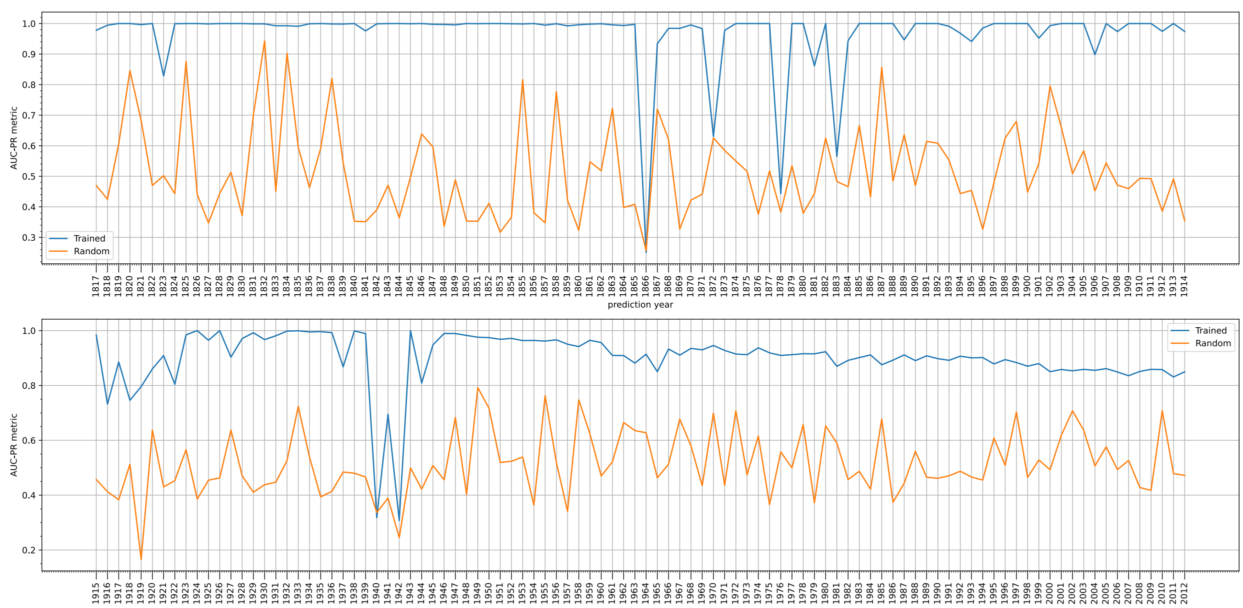


*Figure 7: Line Graph of Model Results*

## e. Model Accuracies

Below is an example of a year – 1943 – in which the model's results were almost entirely accurate. The top chart – depicted using grey lines – is reality, and the bottom chart – depicted using blue lines – is the model's prediction. All edges were correct except for one that used a contemporarily nonexistent node.

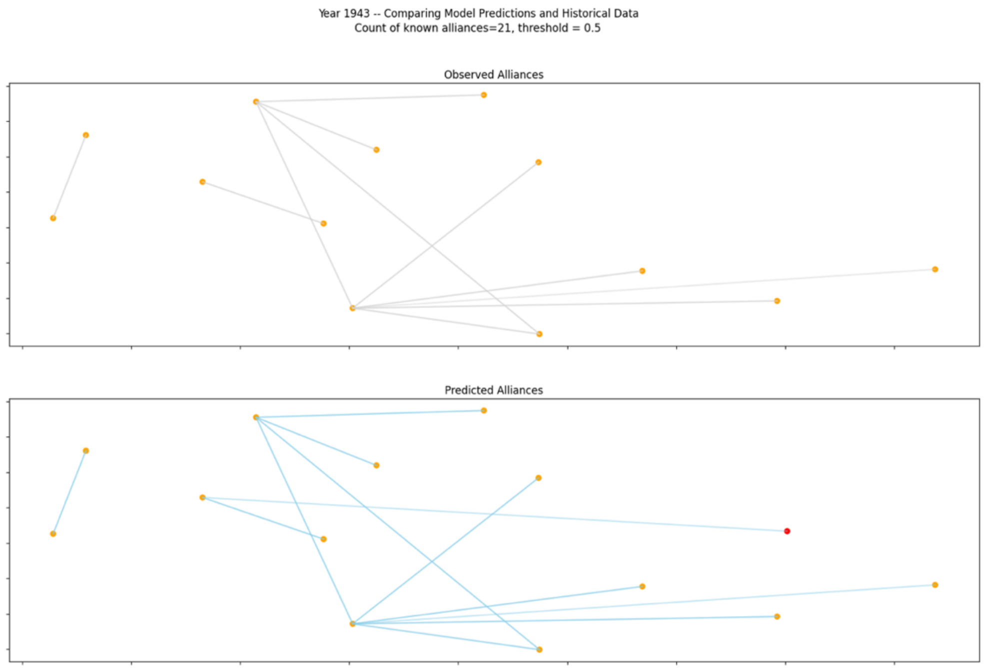


*Figure 8: Reality vs. Model Prediction for 1943*

Three additional examples of accurate years in descending order follows – 1932, 1832, and 1823.

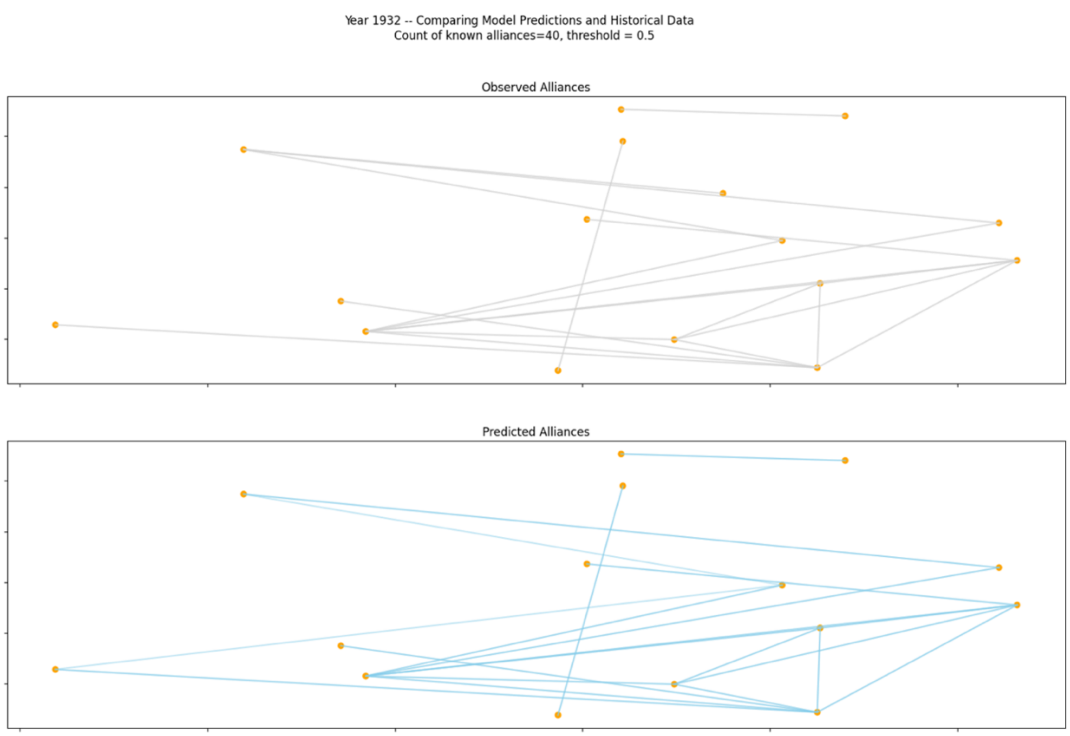


*Figure 9: Reality vs. Model Prediction for 1932*

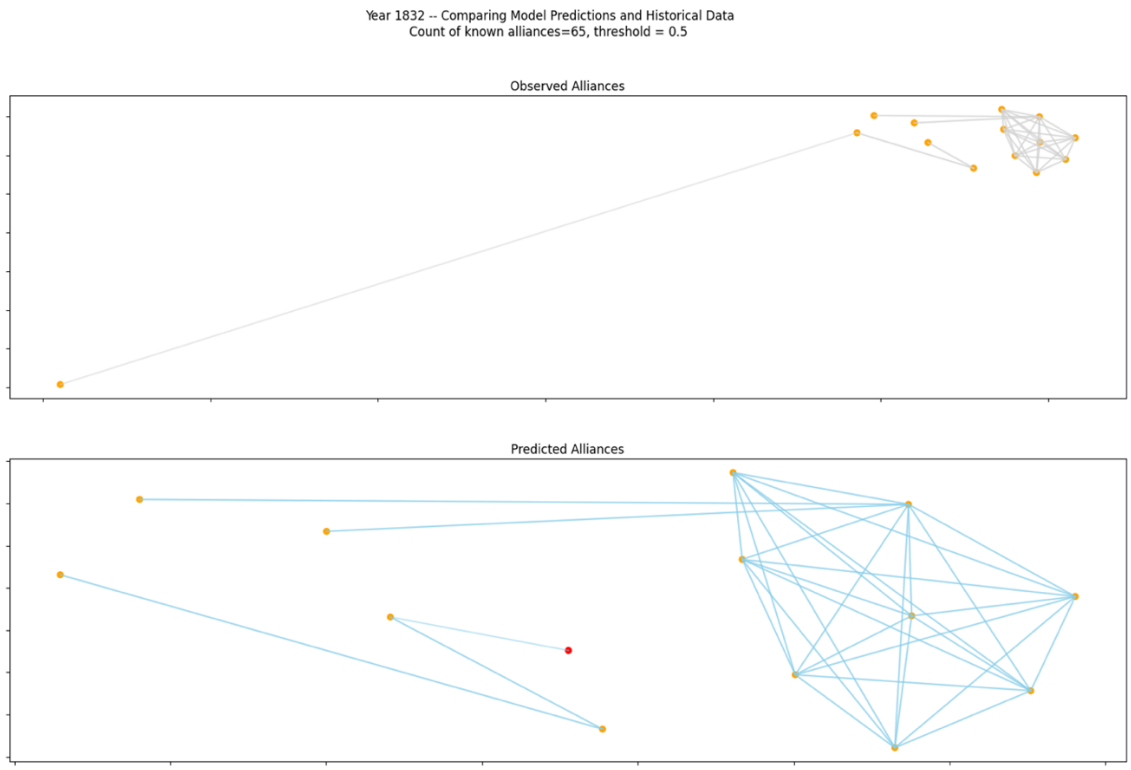


*Figure 10: Reality vs. Model Prediction for 1832*

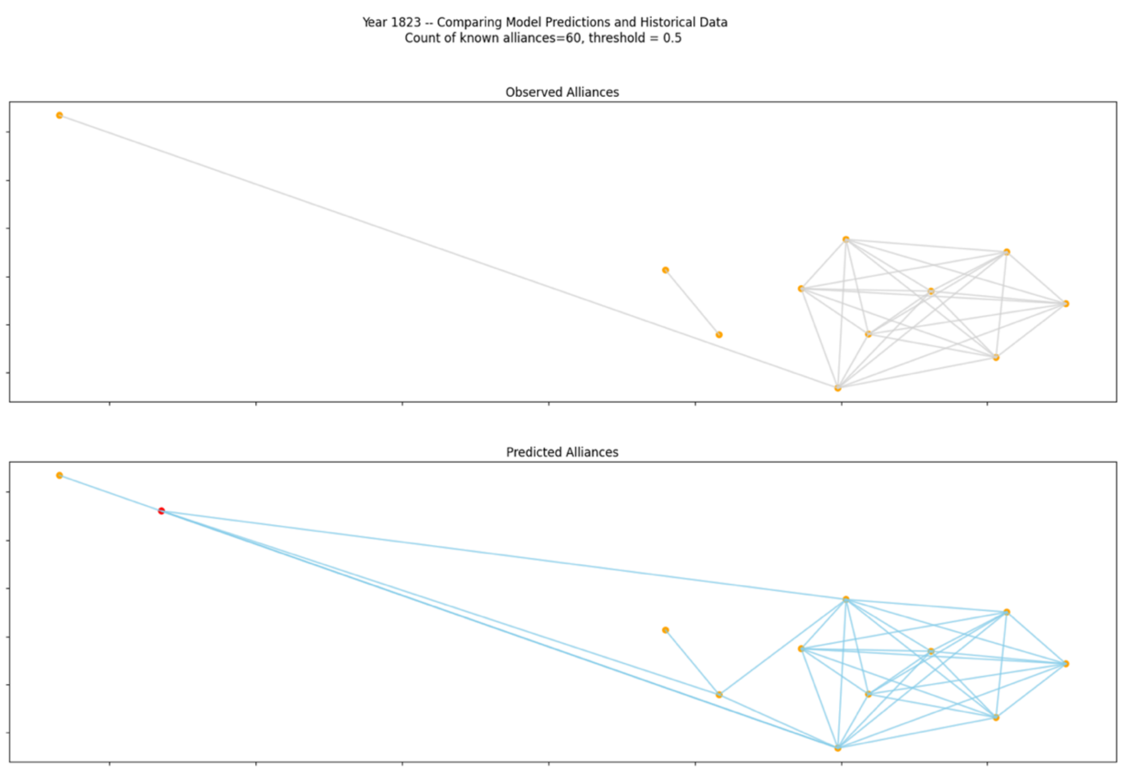


*Figure 11: Reality vs. Model Prediction for 1823*

The model also uses tables to analyze whether its predicted new edge additions (a new alliance formed) and new edge deletions (a newly dissolved alliance) are correct. This is another way to display and analyze the results that omits alliances that carried over from previous years. For example, as illustrated in both the graphs above and the table below, 1823's predicted edge additions were correct.

| | year | threshold | alliance | recent_change | model_prediction | model_score |
|---|---|---|---|---|---|---|
| **0** | 1821 | 0.001 | [Austria-Hungary, Two Sicilies] | added | Correct | 0.033796 |
| **0** | 1821 | 0.001 | [Netherlands, Spain] | added | Correct | 0.039035 |
| **0** | 1823 | 0.001 | [Austria-Hungary, United Kingdom] | added | Correct | 0.005360 |
| **0** | 1823 | 0.001 | [Austria-Hungary, Russia] | added | Correct | 0.006087 |
| **0** | 1823 | 0.001 | [Germany, United Kingdom] | added | Correct | 0.624011 |
| **0** | 1823 | 0.001 | [Russia, United Kingdom] | added | Correct | 0.551497 |
| **0** | 1823 | 0.001 | [Germany, Russia] | added | Correct | 0.351757 |

*Figure 12: Table of Newly Added and Deleted Edges for 1821 and 1823*

### f. Model Inaccuracies

Below is an example of a year – 1940 – in which the model's results were not entirely accurate. 1940's inaccuracies were likely influenced by the model training using WW1's data, which was not a good exemplar of what WWII's alliances would look like. Alliances in WWI were primarily preserved, whereas many alliances in WWII dissolved, explaining why the model predicted the presence of several alliances that did not exist at the time. This is further elaborated in an upcoming section.

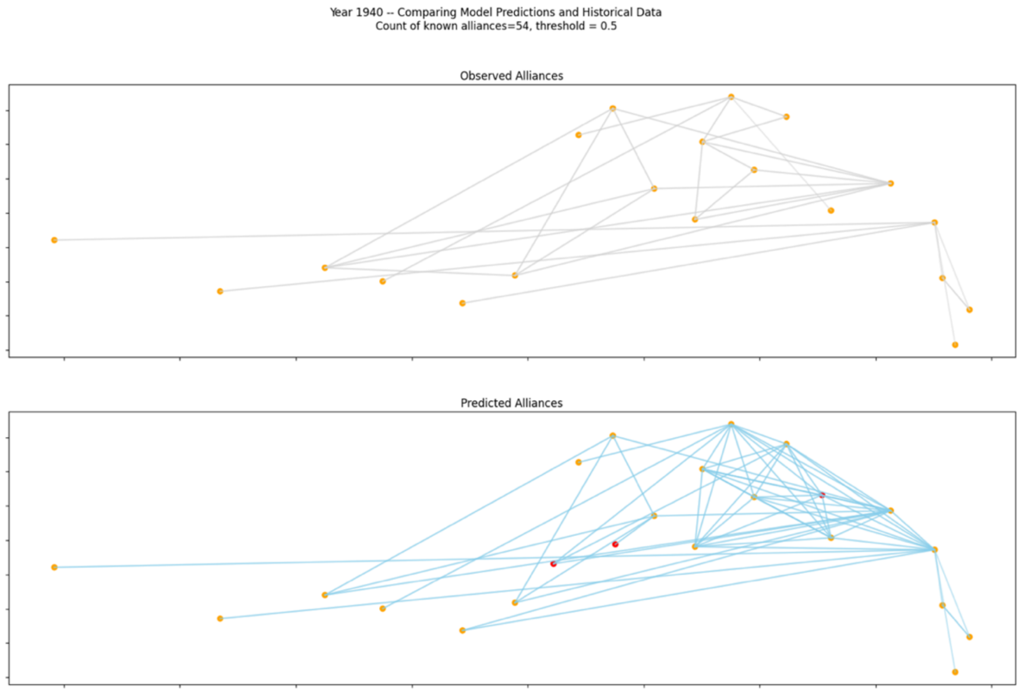


*Figure 13: Reality vs. Model Prediction for 1940*

Three additional examples of less accurate years in descending order follows – 1942, 1883, and 1878.

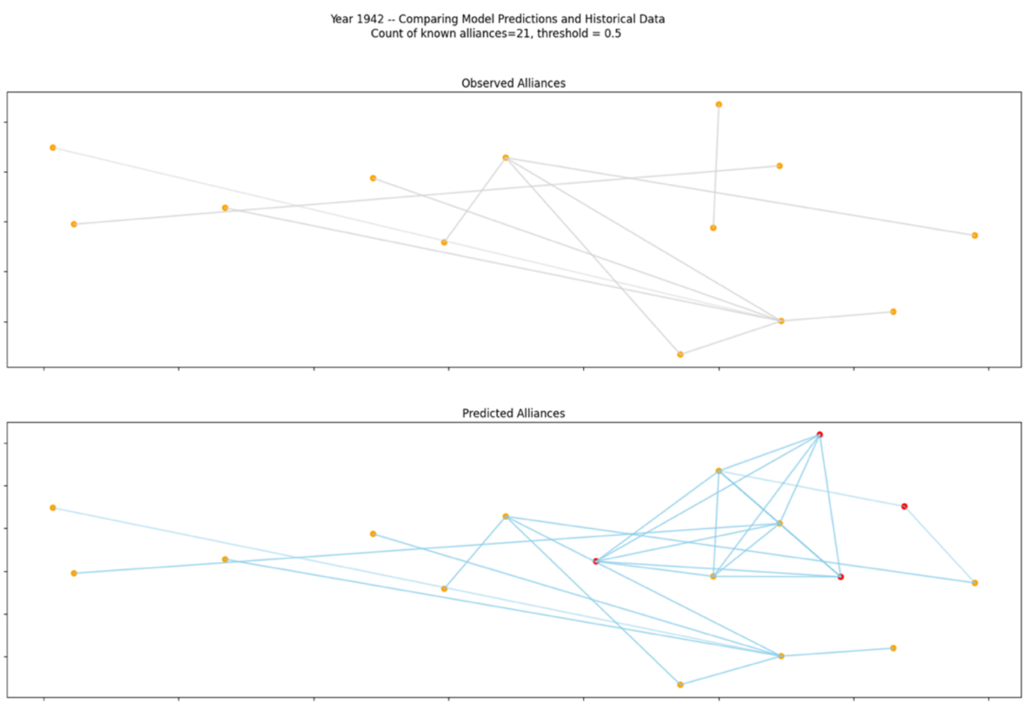


*Figure 14: Reality vs. Model Prediction for 1942*

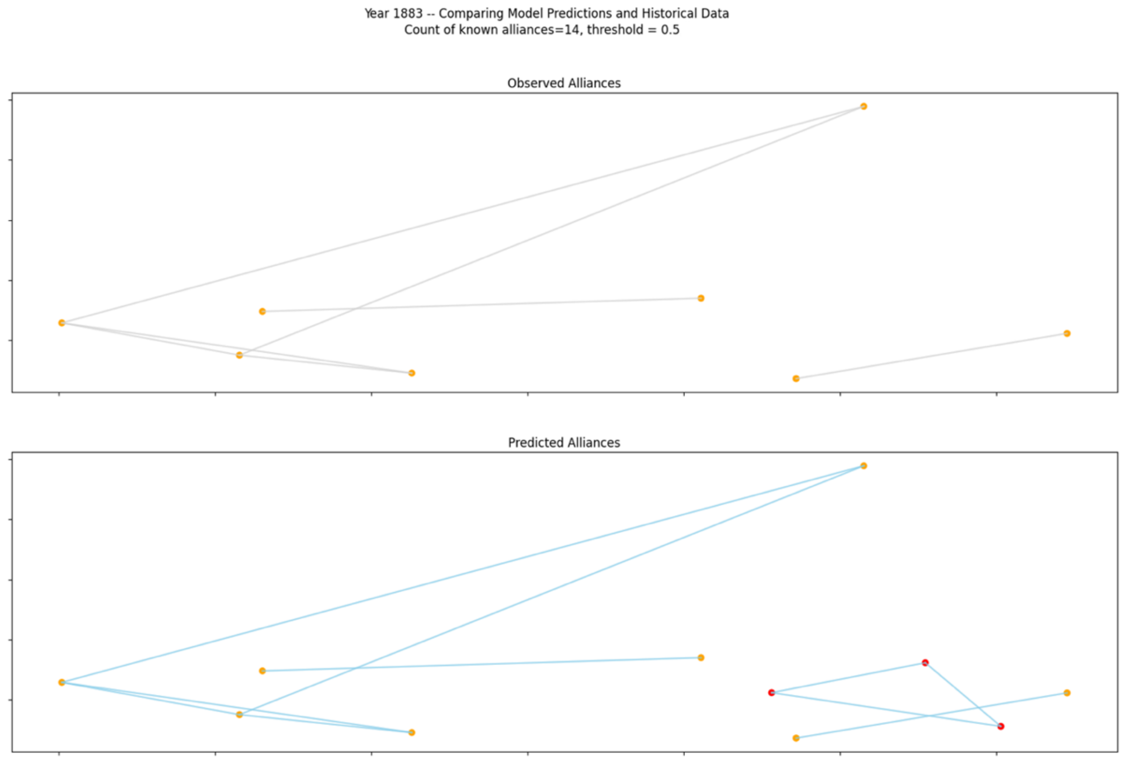


*Figure 15: Reality vs. Model Prediction for 1883*

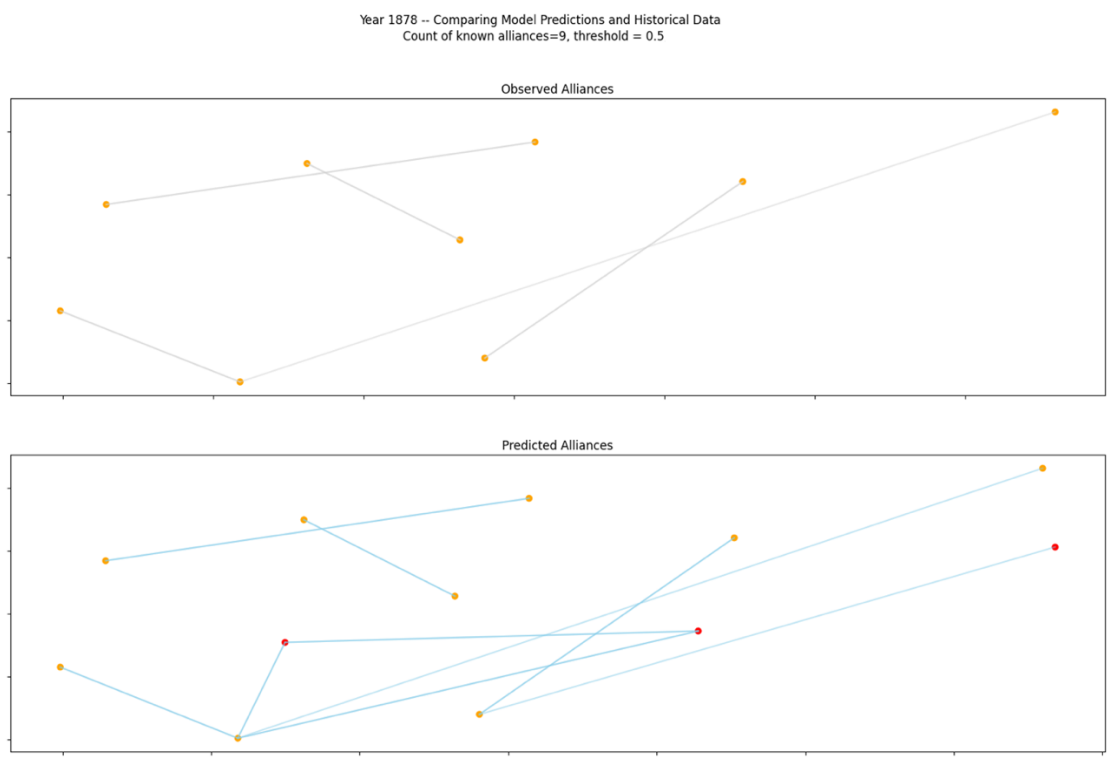


*Figure 16: Reality vs. Model Prediction for 1878*

Analogous to the table depicting the model's accuracies, below is a table depicting an instance when the model was inaccurate in predicting new alliance formation/dissolution. All the predictions of changes from 2005-2006 were accurate – predictions of alliance dissolutions in Africa – and all the predictions of changes from 2010-2011 were inaccurate – predictions of alliance formations in the Middle East. Consequently, next steps would be to conduct sensitivity analyses to examine if any current node or edge features worse model performance when predicting alliances in the Middle East. Additionally, the modeler could identify a separate node or edge feature that could be added to the model to improve this region's accuracy.

| | year | threshold | alliance | recent_change | model_prediction | model_score |
|---|---|---|---|---|---|---|
| **0** | 2006 | 0.999 | [Congo, Kenya] | deleted | Correct | 0.386309 |
| **0** | 2006 | 0.999 | [Angola, Tanzania] | deleted | Correct | 0.398941 |
| **0** | 2006 | 0.999 | [Angola, Kenya] | deleted | Correct | 0.392653 |
| **0** | 2006 | 0.999 | [Uganda, Zambia] | deleted | Correct | 0.393749 |
| **0** | 2011 | 0.999 | [Syria, Tunisia] | added | Wrong | 0.416494 |
| **0** | 2011 | 0.999 | [Lebanon, Syria] | added | Wrong | 0.417023 |
| **0** | 2011 | 0.999 | [Iraq, Syria] | added | Wrong | 0.194316 |
| **0** | 2011 | 0.999 | [Jordan, Syria] | added | Wrong | 0.258273 |
| **0** | 2011 | 0.999 | [Syria, United Arab Emirates] | added | Wrong | 0.256606 |
| **0** | 2011 | 0.999 | [Saudi Arabia, Syria] | added | Wrong | 0.260151 |

*Figure 17: Table of Newly Added and Deleted Edges for 2006 and 2011*

## g. Notable Dips in Accuracy

Here is a scatter plot depicting the results of the trained model in blue, compared to the results of a random model in orange. Overall, the trained model performed substantially better than the random model, but there are ten years that did not reach a 0.8 AUC-PR.

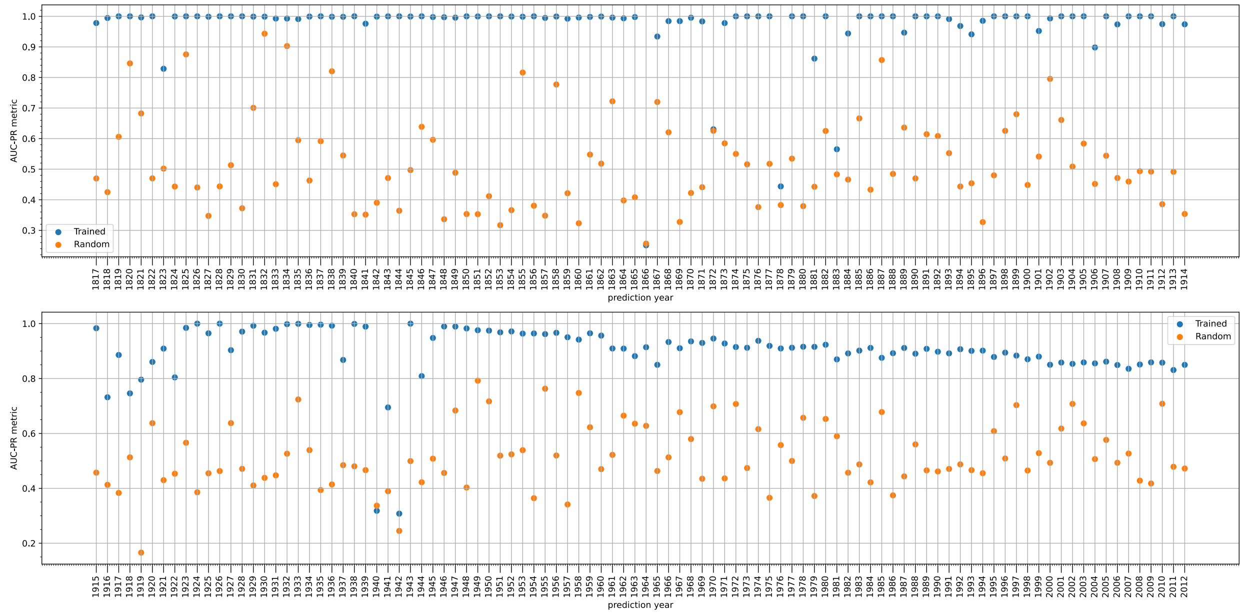


*Figure 18: Scatter Plot of Model Results*

The model dipped in accuracy to below a 0.8 AUC-PR for ten years. Here is a preliminary analysis of events that likely caused those temporary deteriorations:

In 1866, the **Austro-Prussian War** began and ended. Also known as the Seven Weeks' War, this conflict between Austria and Prussia resulted in the dissolution of the German Confederation and the formation of the North German Confederation under Prussian leadership, which likely affected the model's accuracy due to new node emergence/disappearance.[31] Simultaneously, the **Third Italian War of Independence** took place, in which Italy allied with Prussia against Austria, leading to the annexation of Venetia by Italy.[32] In preparation for both the Austro-Prussian War and the Third Italian War of Independence, earlier in the year, Italy and Prussia formed the **Italo-Prussian Alliance** against Austria, which also likely impacted the model's edge prediction accuracy because there was no historical precedent for this cooperation.[33]

In 1872, the **League of the Three Emperors** was formed between Austria-Hungary, Germany, and Russia to preserve the social order of the conservative powers of Europe and maintain peace between Austria-Hungary and Russia.[34] However, the Three Emperors' League was not a formal defense alliance and was instead an informal agreement between these three countries' emperors; hence it was not part of the model data, which only recorded formal defense alliances.[35]

In 1878, the **League of the Three Emperors** was dissolved due to conflicts of interest in the Balkans between Austria-Hungary and Russia.[36] In its place, Austria-Hungary and Germany formed the **Dual Alliance**, promising mutual support against possible Russian aggression.[37] If the model had been trained using data on informal agreements like the League of the Three Emperors, its dissolution would likely have hinted at a deterioration in relations between Russia and Austria-Hungary/Germany. Additionally, the **Treaty of San Stefano** was signed, ending the Russo-Turkish War and granting independence from Ottoman rule to Bulgaria.[38] Later that year, the **Congress of Berlin** was held to revise the Treaty of San Stefano, resulting in the **Treaty of Berlin**, which redrew the map of the Balkans and recognized the independence of Romania, Serbia, and Montenegro, which likely impacted the model's accuracy due to new node emergence/disappearance.[39]

In 1883, Germany, Austria-Hungary, and Italy further developed their 1882 **Triple Alliance**.[40] This was a defensive military alliance promising mutual support in the event of an attack by any other great power and would eventually play a significant role in intensifying World War I.

In 1916, World War I was intensifying, and **Romania** entered the war on the side of the Allies, aligning with countries like France, Italy, Japan, Russia, and the United Kingdom. Romania joined the Allies primarily due to domestic political goals, specifically wanting to acquire the territory of Transylvania.[41] This likely impacted the model's accuracy because the model did not include features such as domestic political agendas.

In 1918, several wartime events occurred. First, several countries declared independence, including **Lithuania**, **Estonia**, and **Latvia**, which likely impacted the model's accuracy due to new node emergence/disappearance.[42] Second, the **Treaty of Brest-Litovsk** was signed between Russia and the Central Powers, ending Russia's involvement in World War I, due to domestic turmoil during the Russian Revolution and the rise of the Bolsheviks.[43] This likely impacted the model's accuracy because the model did not train with a domestic political agenda dataset. Third, the alliance of the **Central Powers** (Germany, Austria-Hungary, the Ottoman Empire, and Bulgaria) effectively dissolved with the **end of World War I** and the signing of the armistice on November 11, which also likely impacted the model's edge prediction accuracy because there was no historical precedent for this dissolution.[44]

In 1919, the **Treaty of Versailles** was signed – which imposed military limitations on Germany and effectively isolated it from forming or maintaining defense alliances – impacting the model's edge prediction accuracy because, again, there was no historical precedent pre-WWI.[45] World War I was an unprecedented global event, and, because there were no preceding similar events, the model had difficulty in predicting that military alliances would steadily increase throughout the war, as illustrated by Figure 6.

Consequently, the model's performance fluctuated, performing well in 1914 and 1915, less so in 1916, better in 1917, and worse in 1918 and 1919.

In 1940, **World War II** was intensifying. In 1941, the **Attack on Pearl Harbor** brought the United States into the war, forging new wartime defense alliances.[46] This likely impacted the model's accuracy because the model did not train with a dataset that detailed notable attacks on neutral states. Additionally, the model's lower accuracy levels in 1940, 1941, and 1942 were likely exacerbated by the presence of only one comparable historical precedent: WWI. During WWI, alliances had steadily increased, which the model learned from. However, during WWII, alliances drastically decreased until near the end of the war, as illustrated by the graph above. Consequently, because the model had only had one analogous prior event – WWI – to learn from and WWII deviated from WWI's pattern, the model's accuracy levels temporarily decreased. However, going forward, the model will likely benefit from training using two extreme events located on opposite ends of the alliance pattern spectrum.

## h. Future Features

The above temporal analysis illuminated what additional features should be added to the model to mitigate its inaccuracies in those ten years.

### *i. Domestic Features*

Several alliance formations and dissolutions that took place in those ten years were due to domestic factors. First, the model should train using a data set that includes domestic turmoil data such as Russia leaving WWI during its October Revolution and Ireland's independence movement. Second, the model should use an ongoing data set that tracks domestic political agendas (e.g. Romania's domestic political agenda regarding Transylvania during WWI). Third, the model should incorporate data sets that feature domestic ideological trends like nationalism, communism, liberalism, imperialism. Fourth, the model would benefit from using a leadership ideology dataset (e.g. Prussian strategy under Otto von Bismarck's leadership). Finally, the model would likely benefit from training using a data set on domestic influence networks, as power brokers within governments and military hierarchies play pivotal roles in alliance dynamics. However, this information may be difficult to find through purely open-source investigations.

### *ii. Economic Features*

Additional economic features would also be helpful for modern predictions. A prospective feature is interstate trade, as some scholars theorize that countries are more likely to align on defense matters with strong trading partners. A data set tracking natural resources like oil would also likely be important to improving model accuracy, particularly in regions like the Middle East. Finally, information on sanctions or financial pressures would likely also be advantageous, as these often destabilize alliances or force concessions.

### *iii. International Features*

The model is based on official, interstate defense alliances. However, training using open-ended agreements, resolution frameworks, and treaties would likely improve the model's accuracy, as these international bilateral and multilateral instruments are indicative of strengthening and weakening relationships.

A database of these instruments would likely include overarching war-ending agreements. For example, the Moscow Peace Treaty ended the Winter War between Finland and the USSR and heightened concerns among neighboring countries, influencing their defense alliances in future years. Another example is the 1878 Treaty of Berlin, whose territorial stipulations frustrated several Balkan nations – triggering nationalist movements and future alliances – and Russia – weakening Russia's ties with Germany and Austria-Hungary and strengthening Germany and Austria-Hungary's relationship. This could also be expanded to include not just war-ending agreements but also wartime agreements like the Declaration by

United Nations, which was a WWII agreement among 26 Allied nations committing them to cooperate against the Axis Powers while omitting mutual defense obligations.

The database should also include diplomatic arrangements like the League of the Three Emperors (which established reciprocal neutrality) and multilateral organizations like the League of Nations (which established a collective security approach). A dynamic list of frameworks is also key to improving the model's accuracy. For example, the 1916 secret Sykes-Picot Agreement between Britain and France planned the anticipated post-WWI division of the Ottoman Empire's Middle East territories. Because the Sykes-Picot agreement was a framework rather than a defense alliance, the model did not include its anticipated stipulations, which likely impacted the model's accuracy due to new node emergence/disappearance. Similarly, the 1941 Atlantic Charter outlined a collaborative United States/United Kingdom vision for the post-WWII world, and its stipulations would likely have helped the model predict the years immediately following WWII.

The database should also contain a dynamic list of bilateral and multilateral meetings between country leaders. For example, the March 2021 Alaska meeting between American and Chinese leaders signaled a deterioration in US-China relations, and the February 2025 Trump-Zelensky Oval Office meeting signaled a deterioration in US-Ukraine relations. Finally, the model should also train on unilateral foreign policy principles. For instance, the 1980 Carter Doctrine, a notable example of US foreign policy, committed the US to militarily defending the peace in the Persian Gulf, deepening US security ties and alliances in the Middle East.

***iv. Military Features***

Several military features would be indicative of evolving alliance trends. First, the current model has trained using a list of countries involved in each major interstate war, and this should be expanded to include the "winners" and "losers" of each war. Second, a database containing surprise military actions like attacks on neutral states, as these typically influenced that neutral state's future alliances. Notable examples include the 1941 Japanese attack on Pearl Harbor that drew the neutral US into WWII; the 1939 Soviet invasion of neutral Finland that began the Winter War; the 1914 German invasion of neutral Belgium that drew Belgium into WWI; and the 1940 German invasion of neutral Norway that drew Norway into WWII. Third, the model would likely benefit from using comparative military power as feature, as analyzing the relative strengths of nations involved in various alliances would help signal major shifts during global events like WWI and WWII. However, it is important to note that comparative military power is difficult to quantify, particularly when limited to open-source information. Fourth, military mobilization data would likely improve the model's accuracy, as this comprises military preparedness, troop movements, and war plans, which help signal alliance shifts. However, this data likely does not exist for past historical wars, so this feature might be more helpful in predicting future modern wars. Additionally, much of this information is likely limited to classified satellite data rather than open-source information.

### i. Sensitivity Tests

Sensitivity analyses evaluate how changes in input variables influence the model's output and overall accuracy. They help modelers discern the relative importance of individual parameters (i.e. the node and edge features). In other words, this project's sensitivity analyses identified the most and least important features, thereby helping to simplify the model, improve efficiency, and provide policy insights on which areas are most impactful for defense alliances (e.g. religion vs. regime type).

Sensitivity analyses work by comparing the latest year's results (i.e. 2012's graphs) with previous years' results. First, the study analyzed the proximity feature: an edge feature that depicted the distance between states. In addition to comparing the proximity feature with the other three features – religion, regime type, and interstate wars – (which is detailed later in this section), a sensitivity analysis manipulated the proximity feature to analyze Kautilya's Mandala Theory. Initially, the study used traditional data analysis to examine the theory. The theory states that a country is more likely to be enemies with its direct neighbors and defense allies with the states that are on the opposite side of their enemies.[47] In other words, states are more likely to be enemies with secondary states and allies with tertiary states. The four graphs below visualize defense alliance counts by year and proximity type – with blue depicting defense alliances between direct neighbors/secondary states and orange depicting defense alliances between tertiary states – from 1816–2012. As illustrated, the count of alliances between tertiary states is much higher than the count of alliances between secondary states, showing that states are much more likely to be allies with tertiary states than secondary states, thereby helping substantiate Kautilya's Mandala Theory. It is worth noting that the strength of this theory wanes with time when transitioning from earlier years to later years, although it is still valid.

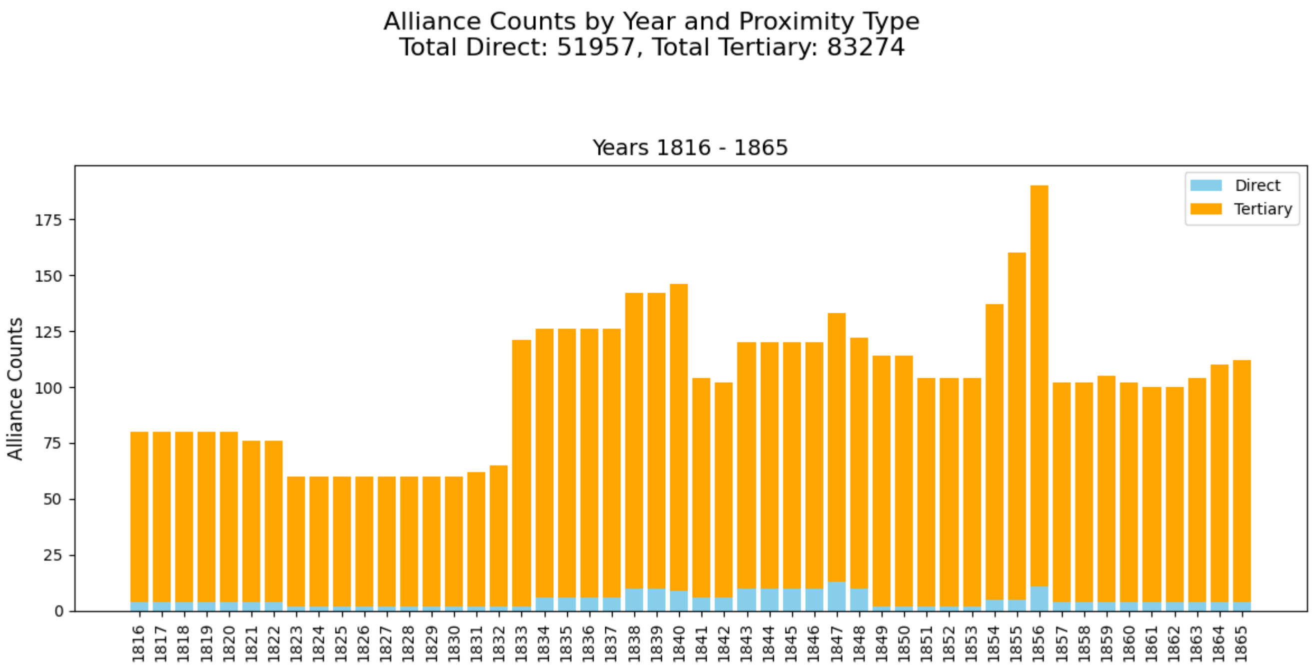


*Figure 19: Bar Chart of Secondary and Tertiary Interstate Relationships from 1816-1865*

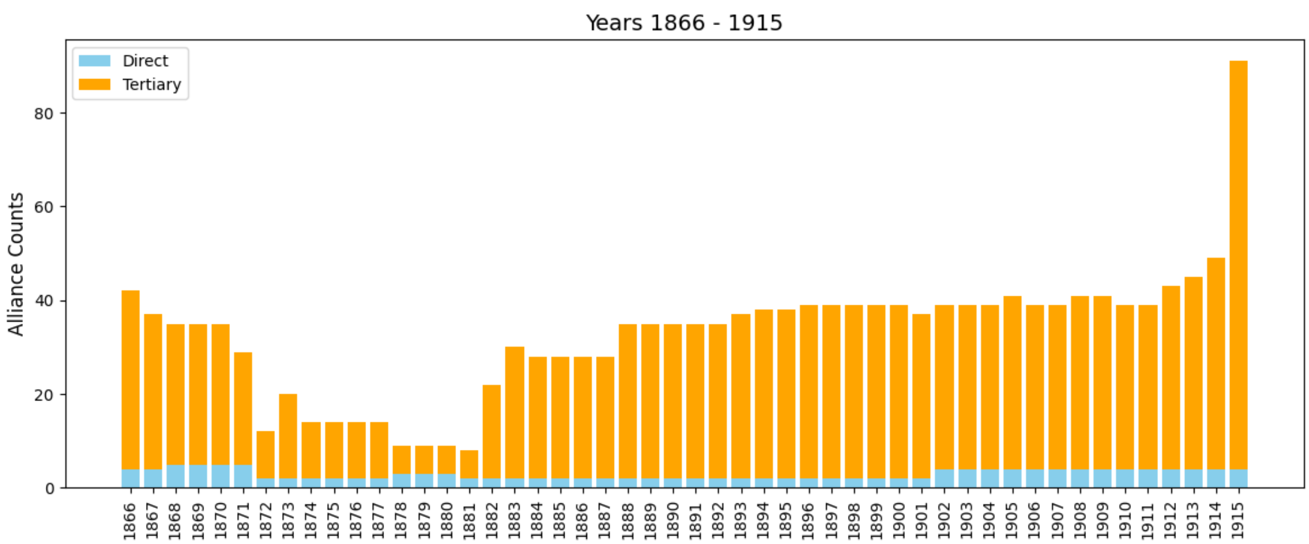


*Figure 20: Bar Chart of Secondary and Tertiary Interstate Relationships from 1866-1915*

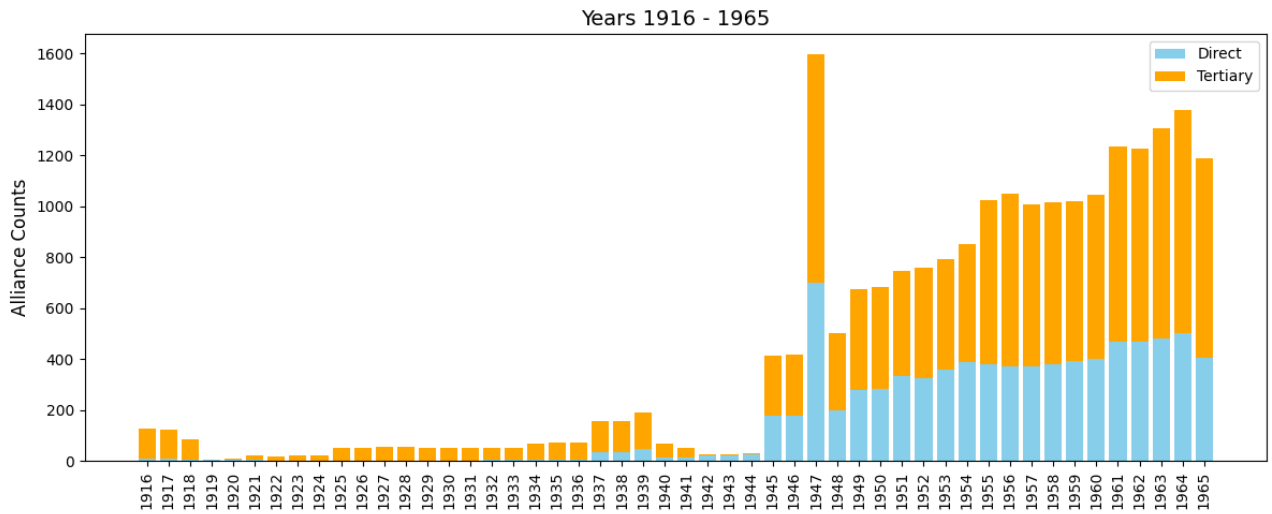


*Figure 21: Bar Chart of Secondary and Tertiary Interstate Relationships from 1916-1965*

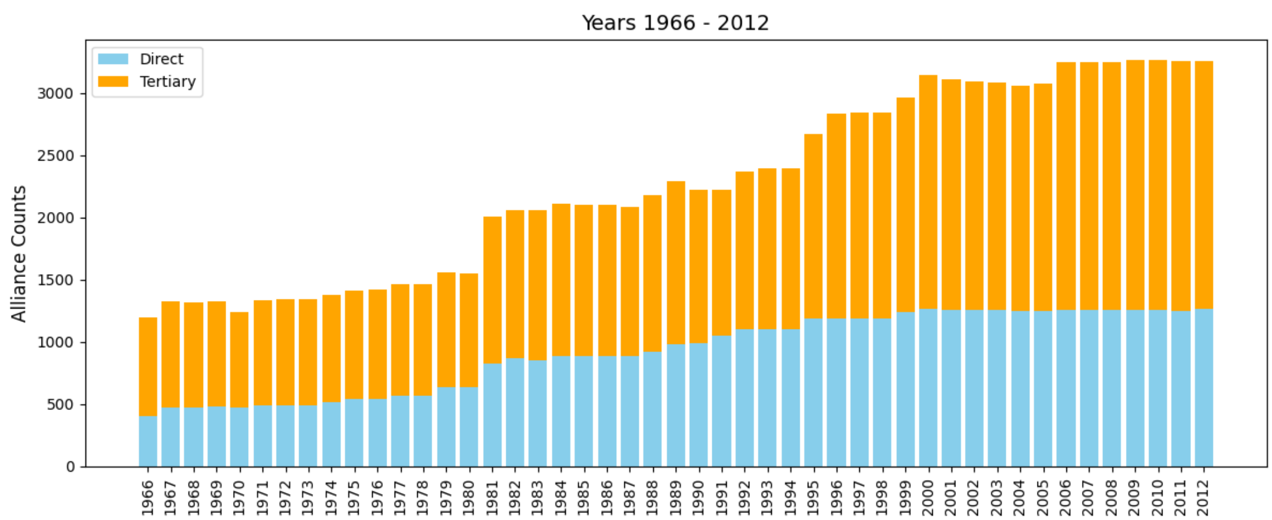


*Figure 22: Bar Chart of Secondary and Tertiary Interstate Relationships from 1916-2012*

Next, the study applied sensitivity analysis to further help discern if secondary relationships are less important to predicting the presence of a defense alliance. Note that typically ratios of 1 yield better performances for this model compared to ratios of 10, except for years that exhibited dips in accuracy such as 1823. For reference, the proximity feature data used to train the model only includes secondary relationships and omits tertiary relationships, so the below sensitivity analysis is limited to secondary relationships. As anticipated, the model outputs for 2012 and 1823 with the proximity feature (i.e. secondary relationship data) included and excluded did not substantially impact the accuracy. The full results are depicted below in Figure 23.

| Negative Ratio | Year | Feature Comparison | Threshold | Accuracy | Precision | Recall | F1 Score | AUC-PR |
|---|---|---|---|---|---|---|---|---|
| 1 | 2012 | With Direct Proximity | 0.343033 | 0.776021 | 0.706772 | 0.944225 | 0.808423 | 0.836045 |
| 1 | 2012 | Without Direct Proximity | 0.318332 | 0.780413 | 0.699065 | 0.985507 | 0.817933 | 0.836823 |
| 1 | 1823 | With Direct Proximity | 0.983146 | 0.923077 | 0.892308 | 0.966667 | 0.928 | 0.842915 |
| 1 | 1823 | Without Direct Proximity | 0.927138 | 0.915385 | 0.857143 | 1.0 | 0.923077 | 0.84335 |
| 10 | 2012 | With Direct Proximity | 0.41111 | 0.847127 | 0.199302 | 0.225736 | 0.211697 | 0.18203 |
| 10 | 2012 | Without Direct Proximity | 0.411707 | 0.867928 | 0.270499 | 0.266579 | 0.268525 | 0.211098 |
| 10 | 1823 | With Direct Proximity | 0.520644 | 0.992537 | 1.0 | 0.933333 | 0.965517 | 0.972092 |
| 10 | 1823 | Without Direct Proximity | 0.419301 | 0.980597 | 0.875 | 0.933333 | 0.903226 | 0.957568 |
| 1 | 2012 | With Regime Type | 0.318332 | 0.780413 | 0.699065 | 0.985507 | 0.817933 | 0.836823 |
| 1 | 2012 | Without Regime Type | 0.26329 | 0.76592 | 0.687153 | 0.977163 | 0.80689 | 0.854152 |
| 1 | 1823 | With Regime Type | 0.927138 | 0.915385 | 0.857143 | 1.0 | 0.923077 | 0.84335 |
| 1 | 1823 | Without Regime Type | 0.974293 | 0.915385 | 0.857143 | 1.0 | 0.923077 | 0.813576 |
| 10 | 2012 | With Regime Type | 0.411707 | 0.867928 | 0.270499 | 0.266579 | 0.268525 | 0.211098 |
| 10 | 2012 | Without Regime Type | 0.290605 | 0.769653 | 0.695017 | 0.961792 | 0.806927 | 0.850539 |
| 10 | 1823 | With Regime Type | 0.419301 | 0.980597 | 0.875 | 0.933333 | 0.903226 | 0.957568 |
| 10 | 1823 | Without Regime Type | 0.993429 | 0.923077 | 0.918033 | 0.933333 | 0.92562 | 0.847917 |
| 1 | 2012 | With Religion | 0.318332 | 0.780413 | 0.699065 | 0.985507 | 0.817933 | 0.836823 |
| 1 | 2012 | Without Religion | 0.220988 | 0.757795 | 0.680381 | 0.97321 | 0.800867 | 0.853275 |
| 1 | 1823 | With Religion | 0.927138 | 0.915385 | 0.857143 | 1.0 | 0.923077 | 0.84335 |
| 1 | 1823 | Without Religion | 0.977831 | 0.930769 | 0.882353 | 1.0 | 0.9375 | 0.896262 |
| 10 | 2012 | With Religion | 0.411707 | 0.867928 | 0.270499 | 0.266579 | 0.268525 | 0.211098 |
| 10 | 2012 | Without Religion | 0.330129 | 0.771849 | 0.69648 | 0.964427 | 0.80884 | 0.850675 |
| 10 | 1823 | With Religion | 0.419301 | 0.980597 | 0.875 | 0.933333 | 0.903226 | 0.957568 |
| 10 | 1823 | Without Religion | 0.963572 | 0.953846 | 0.923077 | 1.0 | 0.96 | 0.862035 |
| 1 | 2012 | With Interstate Wars | 0.318332 | 0.780413 | 0.699065 | 0.985507 | 0.817933 | 0.836823 |
| 1 | 2012 | Without Interstate Wars | 0.337552 | 0.776899 | 0.698054 | 0.976724 | 0.814205 | 0.851156 |
| 1 | 1823 | With Interstate Wars | 0.927138 | 0.915385 | 0.857143 | 1.0 | 0.923077 | 0.84335 |
| 1 | 1823 | Without Interstate Wars | 0.995388 | 0.930769 | 0.933333 | 0.933333 | 0.933333 | 0.92959 |
| 10 | 2012 | With Interstate Wars | 0.411707 | 0.867928 | 0.270499 | 0.266579 | 0.268525 | 0.211098 |
| 10 | 2012 | Without Interstate Wars | 0.43397 | 0.785024 | 0.714569 | 0.949934 | 0.815611 | 0.854082 |
| 10 | 1823 | With Interstate Wars | 0.419301 | 0.980597 | 0.875 | 0.933333 | 0.903226 | 0.957568 |
| 10 | 1823 | Without Interstate Wars | 0.986691 | 0.930769 | 0.893939 | 0.983333 | 0.936508 | 0.866239 |

*Figure 23: Sensitivity Analyses Results for Negative Ratios 1 and 10, Organized by Feature*

2012's accuracy results with a negative ratio of 1 are 0.836045 and 0.836823 with and without secondary relationship proximity data, respectively, exhibiting a difference of 0.000777 in accuracy. 1823's accuracy results with a negative ratio of 1 are 0.842915 and 0.84335 with and without secondary relationship proximity data, respectively, exhibiting a difference of 0.000435 in accuracy. 2012's accuracy results with a negative ratio of 10 are 0.18203 and 0.211098 with and without secondary relationship proximity data, respectively, exhibiting a difference of 0.029068 in accuracy. 1823's accuracy results with a negative ratio of 10 are 0.972092 and 0.957568 with and without secondary relationship proximity data, respectively, exhibiting a difference of 0.014524 in accuracy. The difference in accuracy for all four cases is negligible – in the ten-thousandths place for a negative ratio of 1 and in the hundredths place for a negative ratio of 10. Consequently, it is acceptable to assume that secondary relationships (i.e. the current proximity feature) are not as relevant or indicative of forming a defense alliance as other features, which further substantiates Kautilya's theory. A recommendation here would be to add data on interstate tertiary relationships to the model and conduct a similar sensitivity analysis.

The study then applied sensitivity analysis to the other three features, starting with regime type. The model outputs for 2012 and 1823 with the regime type feature included and excluded are depicted in Figure 23. 2012's accuracy results with a negative ratio of 1 increased from 0.836823 with regime data to 0.854152 without regime data. 1823's accuracy results with a negative ratio of 1 decreased from 0.84335 with regime data to 0.813576 without regime data. 2012's accuracy results with a negative ratio of 10 increased from 0.211098 with regime data to 0.850539 without regime data. 1823's accuracy results with a negative ratio of 10 decreased from 0.957568 with regime data to 0.847917 without regime data. It seems that regime type data improves the model's performance in earlier years and worsens the model's performance in modern years.

It would be instructive to determine the period in history when the "without regime" model's performance overtakes the "with regime" model's performance. That would enable further analysis and a determination on how best to incorporate this feature for model training.

Next, the study assessed the religion feature. The model outputs for 2012 and 1823 with the religion feature included and excluded are depicted in Figure 23. 2012's accuracy results with a negative ratio of 1 increased from 0.836823 with religion data to 0.853275 without religion data. 1823's accuracy results with a negative ratio of 1 increased from 0.84335 with religion data to 0.896262 without religion data. 2012's accuracy results with a negative ratio of 10 increased from 0.211098 with religion data to 0.850675 without religion data. 1823's accuracy results with a negative ratio of 10 decreased from 0.957568 with religion data to 0.862035 without religion data. Consequently, religion as an overall model feature is relatively detrimental to the model's performance compared to other features. Training without religion is more helpful for eliminating false positives because the presence of religion causes the model to predict many more negative edges than positive edges for later modern years. In other words, the relative effect of religion with respect to this false positive effect is much more pronounced in modern times. Therefore, re-training the entire model without the religion feature is likely to help eliminate false positives and improve the model's overall accuracy.

Finally, the study analyzed the interstate war feature. The model outputs for 2012 and 1823 with the war feature included and excluded are depicted in Figure 23. 2012's accuracy results with a negative ratio of 1 increased from 0.836823 with war data to 0.851156 without war data. 1823's accuracy results with a negative ratio of 1 increased from 0.84335 with war data to

0.92959 without war data. 2012's accuracy results with a negative ratio of 10 increased from 0.211098 with war data to 0.854082 without war data. 1823's accuracy results with a negative ratio of 10 decreased from 0.957568 with war data to 0.866239 without war data. The interstate war feature seems to behave similarly to the religion feature (re-training the entire model without the war feature would confirm this behavior). Therefore, war as a feature seems overall detrimental to the model's performance compared to the proximity and regime features. Training without the interstate war data is more helpful for eliminating false positives because the presence of this feature causes the model to predict many more negative edges than positive edges for later modern years. In other words, the relative effect of the war feature with respect to this false positive effect is much more pronounced in modern times. Therefore, re-training the entire model without the war feature is likely to help eliminate false positives and improve the model's overall accuracy.

In sum, the proximity/secondary relationships feature does not substantially affect the model's accuracy. The regime feature improves the model's performance in earlier years and worsens the model's performance in modern years. The religion and interstate war features appear overall detrimental to model performance, but their regional performances have yet to be examined given the project's scope and time constraints.

## V: Summary and Actions

### a. Findings

The model discovered important relationships and patterns in the data. Analyses yielded several key findings about how/whether the chosen node and edge features provide insights into the evolution of interstate defense alliances, which enables enhanced conceptual understanding for policymakers.

First, regime type as a node feature improves the model's performance in earlier years and worsens the model's performance in modern years. Second, religion as a node feature and interstate war as an edge feature are both relatively detrimental to the model's performance compared to the other two features.

Third, direct proximity as an edge feature is not substantially significant when improving the model's prediction accuracy. Additionally, this feature's behavior helps substantiate Kautilya's Mandala theory. The theory asserts that a state's direct neighbor (i.e. secondary relationships) is typically its enemy and that the state's neighbor's neighbor (i.e. tertiary relationships) is typically its ally.[48] If the theory is true, secondary relationships as an edge feature are unlikely to be as relevant or indicative of forming a defense alliance as other features. The direct proximity data set exclusively contained secondary relationships, and sensitivity analysis including and excluding this feature exhibited a negligible difference, helping substantiate Kautilya's Mandala theory. Conducting an analogous sensitivity analysis using tertiary relationship data could further verify the theory.

Additionally, there were ten major dips in accuracy levels that were largely attributed to the rise and fall of states:

- 1866: the Austro-Prussian War, Third Italian War of Independence, Italo-Prussian Alliance was formed
- 1872: League of the Three Emperors was formed
- 1878: League of the Three Emperors was dissolved, Dual Alliance was formed, Treaty of San Stefano, Congress of Berlin, Treaty of Berlin
- 1883: 1882 Triple Alliance was further developed
- 1916: Romania joined the Allies in WWI

- 1918: Lithuania, Estonia, and Latvia declared independence, Treaty of Brest-Litovsk, Central Powers Alliance was dissolved, WWI ended
- 1919: Treaty of Versailles
- 1940-1942: WWII intensified, Attack on Pearl Harbor

These discoveries prompted the identification of several potential new features that could mitigate such errors:

- Domestic features such as domestic turmoil, political agendas, ideological trends, leadership ideologies, and internal influence networks.
- Economic features such as interstate trade, natural resources, and sanctions or financial pressures.
- International features such as open-ended agreements, diplomatic arrangements, multilateral organizations, resolutions, frameworks, treaties, leadership meetings, and unilateral foreign policy principles.
- Military features such as wartime "winners" and "losers," surprise military actions, comparative military power, and military mobilization data.

Additional work may be required to corroborate these findings.

### b. Next Steps

The first task is to re-run the entire model with all the different combinations of current node and edge features (with one, two, or three removed) to ensure a more comprehensive verification and validation that spans 1816-2012. In total, this should produce 40 new versions of the model for a comprehensive evaluation.

Next, conduct segmented sensitivity analyses for various purposes. For instance, use the version of the model that omits the regime feature to identify the precise year in which the "without regime" model's performance overtakes the "with regime" model's performance. Then, restrict the regime feature to be used only for the years prior. Another possibility: examine existing node and edge features' performances by region and time-period.

After using sensitivity analyses to identify the detrimental or redundant features, identify and incorporate new features to improve the model's performance. An earlier section listed out potential new domestic, economic, international, and military features. Additionally, as the current proximity feature omitted tertiary relationships in favor of secondary relationships, adding a tertiary relationship feature would likely improve the model's accuracy, per Kautilya's Mandala Theory. Moreover, as model performance on the Middle East seems to be less stable than performance on regions like Africa, identifying some new features that are characteristic of the Middle East – such as oil activity, proxy relationships, and insurgency/terrorist activity – would likely improve the model's accuracy.

After integrating new features into the model and re-training, complete the first and second steps again to re-validate and re-verify the model's performance with more sensitivity analyses. This will ensure the most accurate model possible.

Once satisfied with the model's performance, run the model for a future year (e.g. 2025 through 2030) to assess the model's long-term prediction capability. To do so, assume that the previous years' model-driven extrapolations are accurate and therefore the true static graphs. In other words, to predict 2030's defense alliances, the model should use 1816-2012's online data (as the selected data sets stop at 2012), 2012-2024's researcher-created static graphs, and 2025-2029's model-driven extrapolations or augmented by reasonable assumptions of new data for those years.

### c. Recommendations

Similar research on graph theory – including graph algorithm development and customization, domain use case identification, performance analyses, and data synthesis – should be applied to other national security topics like terrorism networks, illicit trafficking networks, organized crime syndicates, and intelligence networks, depending on the modelers' access to applicable data sets. Researchers could even modify this paper's edge prediction model for edge classification or node prediction problems, depending on computing power.

Additionally, policymakers and policy analysts should use this paper's findings to curb common assumptions. Analysts should not use direct proximity as a sole substantial indicator for potential defense alliance formation. Analysts should also not use seemingly comparable religion and regime type as contemporary indicators of defense alliance formation.

### d. Conclusion

This research focused on two primary strategic insights. First, it successfully identified the relationships between shifts in defense alliances and specific features – direct proximity, regime type, religion, and interstate wars – using a graph theoretic model. Second, it demonstrated that graph theory can accurately predict defense alliance formation and dissolution using temporal data and machine-learned spatial information. The model's high accuracy rates definitively confirm the value of applying this branch of mathematics to derive predictive insights in contemporary security and social science problem sets. Going forward, the historically conspicuous line separating the hard sciences from the social sciences should continue to blur – as illustrated by this project – to skillfully tackle the most complex security problems nation states face today.

# VI: Bibliography

[1] "Formal Alliances (v4.1)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/formal-alliances/.
[2] "COW War Data, 1816 – 2007 (v4.0)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/cow-war/.
[3] "Direct Contiguity (v3.2)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/direct-contiguity/.
[4] "Political Regime." Our World in Data. Accessed January 24, 2025. https://ourworldindata.org/grapher/political-regime.
[5] "World Religion Data (v1.1)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/world-religion-data/.
[6] Chen, Jinyin, Xueke Wang, and Xuanheng Xu. "GC-LSTM: Graph Convolution Embedded LSTM for Dynamic Network Link Prediction." *Applied Intelligence* 52, no. 7 (September 30, 2021): 7513–28. https://doi.org/10.1007/s10489-021-02518-9.
[7] Chen, Jinyin, Xueke Wang, and Xuanheng Xu. "GC-LSTM: Graph Convolution Embedded LSTM for Dynamic Network Link Prediction." *Applied Intelligence* 52, no. 7 (September 30, 2021): 7513–28. https://doi.org/10.1007/s10489-021-02518-9.
[8] Harary, Frank. *Graph Theory*. Addison-Wesley Publishing Company, Inc., 1969.
[9] Harary, Frank, and Robert Z. Norman. *Graph Theory as a Mathematical Model in Social Science*, no. 2 (n.d.). https://www.idiosophy.com/wp-content/uploads/2017/07/harary-norman.pdf.
[10] Balakrishnan, R., and K. Ranganathan. *A Textbook of Graph Theory*. Second Edition. Springer, 2012.
[11] Brigham, Robert C., Ronald D. Dutton, Teresa W. Haynes, and Stephen T. Hedetniemi. "Powerful Alliances in Graphs." *Discrete Mathematics* 309, no. 8 (April 28, 2009): 2140–47. https://doi.org/10.1016/j.disc.2006.10.026.
[12] Lal, Arya. *Alliances in Graph Theory*, 2020. http://117.211.166.25:8082/jspui/bitstream/123456789/303/1/Alliances%20in%20Graph%20Theory.pdf.
[13] Harary, Frank. "On the Notion of Balance of a Signed Graph." *Michigan Mathematical Journal* 2, no. 2 (1953): 143–46. https://doi.org/10.1307/mmj/1028989917.
[14] Harary, Frank, and Jerald A. Kabell. "A Simple Algorithm to Detect Balance in Signed Graphs." *Mathematical Social Sciences* 1, no. 1 (1980): 131–36.
[15] Chiang, Kai-Yang, Cho-Jui Hsieh, Nagarajan Natarajan, Inderjit S. Dhillon, and Ambuj Tewari. "Prediction and Clustering in Signed Networks: A Local to Global Perspective." *Journal of Machine Learning Research* 15 (March 2014): 1177–1213. https://jmlr.org/papers/volume15/chiang14a/chiang14a.pdf.
[16] Doreian, Patrick, and Andrej Mrvar. "Partitioning Signed Social Networks." *Social Networks* 31, no. 1 (January 2009): 247–58. https://www.researchgate.net/publication/248484742_Partitioning_Signed_Social_Networks.
[17] Kumar, Srijan, Francesca Spezzano, V. S. Subrahmanian, and Christos Faloutsos. "Edge Weight Prediction in Weighted Signed Networks." *2016 IEEE 16th International Conference on Data Mining (ICDM)*, December 2016. https://doi.org/10.1109/icdm.2016.0033.

[18] Ahmad, Iftikhar, Muhammad Usman Akhtar, Salma Noor, and Ambreen Shahnaz. "Missing Link Prediction Using Common Neighbor and Centrality Based Parameterized Algorithm." *Scientific Reports* 10, no. 1 (January 15, 2020): 364. https://doi.org/10.1038/s41598-019-57304-y.
[19] Lichtenwalter, Ryan, and Nitesh V. Chawla. "Link Prediction: Fair and Effective Evaluation." *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, August 2012, 376–83. https://doi.org/10.1109/ASONAM.2012.68.
[20] Yang, Yang, Ryan N. Lichtenwalter, and Nitesh V. Chawla. "Evaluating Link Prediction Methods." *Knowledge and Information Systems* 45, no. 3 (October 9, 2014): 751–82. https://doi.org/10.1007/s10115-014-0789-0.
[21] Chen, Jinyin, Xueke Wang, and Xuanheng Xu. "GC-LSTM: Graph Convolution Embedded LSTM for Dynamic Network Link Prediction." *Applied Intelligence* 52, no. 7 (September 30, 2021): 7513–28. https://doi.org/10.1007/s10489-021-02518-9.
[22] Shafique, Khurram Hassan. "Partitioning a Graph in Alliances and Its Application to Data Clustering." University of Central Florida, 2004. https://www.crcv.ucf.edu/papers/theses/Khurram_theses.pdf.
[23] Doreian, Patrick. "A Multiple Indicator Approach to Blockmodeling Signed Networks." *Social Networks* 30, no. 3 (July 2008): 247–58. https://doi.org/doi.org/10.1016/j.socnet.2008.03.005.
[24] Cranmer, Skyler J., Bruce A. Desmarais, and Justin H. Kirkland. "Toward a Network Theory of Alliance Formation." *International Interactions* 38, no. 3 (July 2012): 295–324. https://doi.org/10.1080/03050629.2012.677741.
[25] Doreian, Patrick, and Andrej Mrvar. "Structural Balance and Signed International Relations." *Journal of Social Structure* 16, no. 1 (July 2015). https://doi.org/10.21307/joss-2019-012.
[26] "Formal Alliances (v4.1)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/formal-alliances/.
[27] "COW War Data, 1816 – 2007 (v4.0)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/cow-war/.
[28] "Direct Contiguity (v3.2)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/direct-contiguity/.
[29] "Political Regime." Our World in Data. Accessed January 24, 2025. https://ourworldindata.org/grapher/political-regime.
[30] "World Religion Data (v1.1)." Correlates of War. Accessed January 24, 2025. https://correlatesofwar.org/data-sets/world-religion-data/.
[31] Wawro, Geoffrey. *The Austro-Prussian War: Austria's war with Prussia and Italy in 1866*. Cambridge University Press, 1996.
[32] Barker, Nancy Nichols. "Austria, France, and the Venetian Question, 1861-66." *The Journal of Modern History* 36, no. 2 (June 1964): 145–54. https://doi.org/10.1086/239343.
[33] Comandini, Alfredo. "Italy, Prussia, and Austria, 1866-1916." *Current History* 5, no. 3 (December 1916): 531–33.
[34] Keleher, Edward P., and John Quinn Imholte. "Three Emperors' League." EBSCO, 2023. https://www.ebsco.com/research-starters/politics-and-government/three-emperors-league.
[35] Keleher, Edward P., and John Quinn Imholte. "Three Emperors' League." EBSCO, 2023. https://www.ebsco.com/research-starters/politics-and-government/three-emperors-league.
[36] Keleher, Edward P., and John Quinn Imholte. "Three Emperors' League." EBSCO, 2023. https://www.ebsco.com/research-starters/politics-and-government/three-emperors-league.

[37] Jianzhi, Gao. "The Dual Alliance Under the Perspective of the Theory of Alliance Dilemma: A Study Centered on Macedonian Reforms (1903-1908)." *Bulletin of the Institute of Oriental Studies* 4, no. 1 (February 5, 2024).
[38] Hall, Richard C. "Bulgaria and the Origins of the Balkan Wars, 1912–1913." *War and Nationalism: The Balkan Wars, 1912-1913, and their Sociopolitical Implications*, 2013, 85–99.
[39] Shafer, Kenneth Allen. "The Congress of Berlin of 1878: Its Origins and Consequences." Portland State University, 1989. https://doi.org/10.15760/etd.5811.
[40] Schmitt, Bernadotte E. "Triple Alliance and Triple Entente, 1902-1914." *The American Historical Review* 29, no. 3 (April 1924): 449–73.
[41] Torrey, Glenn E. *Romania's Entry into the First World War: The Problem of Strategy*. Vol. XXVI. Emporia, Kansas: School of Graduate and Professional Studies of the Emporia State University, 1978.
[42] Fletcher, William A. "The British Navy in the Baltic, 1918-1920: Its Contribution to the Independence of the Baltic Nations." *Journal of Baltic Studies* 7, no. 2 (1976): 134–44. https://doi.org/10.1080/01629777600000141.
[43] Chernev, Borislav. "The Brest-Litovsk Moment: Self-Determination Discourse in Eastern Europe before Wilsonianism." *Diplomacy & Statecraft* 22, no. 3 (2011): 369–87. https://doi.org/10.1080/09592296.2011.599635.
[44] Lloyd, Nick. "Ending the War." In *The Routledge History of the First World War*, 1st ed. Routledge, 2024.
[45] Grimmer-Solem, Erik. "The Versailles Treaty and the German Imperial Mindscape: The Navy, the Colonial East, and the Impossible Peace in Postwar Germany." Essay. In *The Making of a World Order*, 1st ed. Routledge, 2023.
[46] Hixson, Walter L. *Pearl Harbor in History and Memory: The American Experience in World War II*. Vol. 4. Routledge, 2003.
[47] Banerjee, Aishi. *The Mandala Theory Revisited: Kautilya's Diplomacy in Contemporary Geopolitical Landscape*, January 31, 2025. https://doi.org/10.2139/ssrn.5054075.
[48] Banerjee, Aishi. *The Mandala Theory Revisited: Kautilya's Diplomacy in Contemporary Geopolitical Landscape*, January 31, 2025. https://doi.org/10.2139/ssrn.5054075.