Building a Reliable, Dynamic and Temporal Synthetic Model of the World Trade Web

Sean Kennedy[1], Michael Wish[2], Philip Smith[3], James Sherrell[3], William Shields[4], and Ralucca Gera[3]

[1] Department of Electrical and Computer Engineering,
Naval Postgraduate School, Monterey, CA
[2] Department of Physics,
Naval Postgraduate School, Monterey, CA
[3] Department of Applied Mathematics,
Naval Postgraduate School, Monterey, CA Email: rgera@nps.edu
[4] Department of Operations Research,
Naval Postgraduate School, Monterey, CA

Abstract. This paper presents an accurate, scalable, time-dependent synthetic network model for the World Trade Web (WTW), whose nodes are the different countries that traded from 1996-2020. Using only an initial distribution of countries' global Gross Domestic Product (GDP) as an input, our synthetic network model initializes weighted undirected edges corresponding to total trade between two countries using the presence of a hidden fitness variable dependent on GDP. The synthetic model simulates the creation and deletion of new and existing trade relationships aligned with real-world data. Our results show that this simulated network continues to faithfully approximate the data from WTW about 20 years after creation within a reasonable degree of accuracy.

1 Introduction and Motivation

Massive global trade disruptions have persisted since the beginning of the COVID-19 pandemic. Impediments to everyday life have been commonplace, and unforeseen shortages in goods have stressed traditional supply chains. Accurate, scalable, and interpretable models are essential for mitigating future disruptions to global supply chains. An accurate approximation of real data through synthetic networks could provide extraordinary results in the form of data for simulations, analysis of correlations between factors, or predictive models. So long as we capture the underlying features of our real-world data, synthetic networks are computationally cheap to produce, easily scalable and may reduce bias when constructed to mimic desired properties of the real data.

Thus, in this research we develop a synthetic network generator that closely approximates the real-world World Trade Web (WTW) using gross domestic product (GDP) as an input. We will then create sample networks and compare our time-series synthetic model against the WTW using key network metrics.

In the future, we envision our model being subject to a series of different attacks or natural changes in attempts to simulate real-world disruptions, including both regional and global events. If this synthetic network model reacts similarly to real-world trade disruptions, a synthetic network model generator could be used in the future to plan for and mitigate disruptions in the WTW.

We use UN data [3] to create a weighted, directed temporal network [9], where each time slice of the network represents a year's worth of imports and exports in billions of US dollars for 190 countries. This temporal network represents world trade data from 1996 to 2020. All traded goods are considered and valued in 2022 US dollars. We also acquired world GDP data from the same timeframe from the World Bank [13].

2 Related Work

Interactions in many technological, biological, and social fields have been modeled as complex networks [11]. Econometrics is an especially interesting field where networks can provide additional explanatory power for both individual behavior and aggregate outcomes [4]. Of specific interest are the properties of individual nodes within a network that affect the probability of forming additional connections [1]. While the traditional "rich get richer" approach from the Barabási-Albert synthetic network model works well for network growth, we seek improved modeling for networks with specific interactions, such as the WTW.

Connecting two vertices in a relatively scale-free environment when the bidirectional edge creates a mutual benefit is a standard interaction model [2]. In this model, each node from the N node choices is assigned some fitness parameter, taken from a distribution $\rho(x)$. That is, for every pair of vertices (i, j), an edge is drawn with probability $f(x_i, x_j)$. If f is constant for all nodes pairs, this will produce a standard Erdős-Rényi Model, but a more dynamic probability function will produce a more scale-free model [5].

These fitness-based models have been successfully used to model the World Trade Web in both directed and undirected network models using unweighted edges [8, 6]. That is, we create a network whose nodes are the countries in a given data set, say N countries, and edges are added by using hidden fitness variable x_i proportional to i^{th} country's GDP: the probability that an edge between countries is added i and j $(1 \leq i \neq j \leq N)$ modeled as

$$P_L[x_i, x_j] = \frac{\alpha_0 x_i x_j}{1 + \beta_0 x_i x_j}, \tag{1}$$

where α_0 and β_0 are free parameters of the model input [6]. From this probability, we can compute the expected number of links, L_{exp} as

$$L_{exp} = \sum_{i \neq j} P_L[x_i, x_j]. \tag{2}$$

This metric is used for comparison between real and synthetic results in Section 4 [8].

Due to the nature of the WTW, it is not necessary to expand a model to the directed case. After the collapse of the USSR, the number of countries has only increased slowly, while the number of trade relationships has increased significantly faster, visible as the WTW network becomes denser and denser. If we were to analyze the WTW in its directed form, we would see mutually directed edges between countries becoming more and more common as time goes on. Eventually, if node i was connected to node j in the directed network, node j would be connected to node i with a high probability. Thus, it becomes feasible to view this directed network as an undirected network, with a low probability p that each individual bidirectional edge be replaced with a directional edge in an arbitrary direction [7]. While there is some loss of information in reduction to a unidirectional network, important trade values can be recovered [8].

In order to maintain a consistent model across years, our network must be a temporal network in which each time slice represents each year of the data set. Many temporal networks have additional constraints, but since the WTW is mainly an industrial and infrastructural network, most dynamic concepts such as latency and efficiencies may probably be safely ignored [9].

3 Methodology

Our methodology for creating time-varying, weighted synthetic WTW networks is based on existing methods for generating static instances of unweighted synthetic WTW networks. Previous research identified that the existence of trade relationships between countries is strongly dependent on the relative sizes of the countries' GDPs [6, 8]). This insight was used to develop a method for synthetic WTW generation that takes in country GDP data as an input, and provides an unweighted, undirected network as an output. The resulting unweighted, undirected network possesses a degree distribution approximating the real-world WTW when the corresponding real-world GDP data is provided as input [6, 8]). Our work extends this method through: (1) random generation of appropriate edge weights representing total annual trade between each connected country pair and (2) implementation of a time-based evolution to the synthetic WTW network in which edges with large trade weights are preferentially maintained, edges with smaller trade weights may disappear, and new edges may be created. We incorporate these additional phenomena into our model to provide coherence between each year's model output, which better matches real-world data. While these phenomena may be ascribable to various complex economic or geopolitical mechanisms (e.g., "Globalization"), the development of our model sets aside such underlying mechanisms in favor of heuristically determined factors based on the trade data itself.

The following is a high-level summary of our iterative model. First, we use provided GDP data to create an initial random network based on the methodology of [6]. Then, we assign weights to each edge in the initial network using a heuristic distribution based on each country's GDP. This step completes the initialization of our WTW synthetic network as the first year's network output. Each subsequent year's output is generated from the previous year's output by (1) randomly adding more weighted edges where none currently exist (as a function of the GDP of the two potential countries to be connected), (2) deleting a random portion of existing trade edges according to a heuristically determined distribution, and (3) adjusting trade values for maintained edges based on changes in the associated countries' GDPs. The details of the model are provided in the subsections below.

3.1 Data Acquisition and Cleaning

We obtained our reference WTW data set from the United Nations (UN) Comtrade Database from 1996 through 2020 [3]. We formed a temporal, undirected network from this data, with each year as a time-step. A country's interaction with another country in a given year was modeled as an edge in that time-step, with an edge weight of the total dollar amount of trade in both directions.

Some data cleaning was necessary. Several countries had no trade for specific years, indicating either years in which that country reported nothing to the UN or errors in UN data compilation. To avoid issues in our data analysis that such errors would bring forth, we removed all countries whose trade disappeared for at least one year from the data set. Combined, these countries amounted to a small proportion of our data set, and thus their removal is unlikely to affect the underlying traits of the data.

We also gathered GDP data from the World Bank website [13]. Like the UN trade data, GDP data was not always consistent, with several countries lacking GDP data in certain years. These countries with inconsistent GDP data were likewise removed from our analysis. In total, there were 162 countries that had consistent GDP and trade data for the years 1996 through 2020. It is from this data set that we determined the global parameters for our time-varying synthetic WTW model.

3.2 Synthetic Network: Initialization

We initialize the synthetic WTW model using the method of [6]: taking w_i $(i = 1, \ldots, N)$, as the GDP of the i-th country in the initial year, we first normalize each country's GDP by the mean GDP to obtain each country's so-called "fitness score" x_i as

$$x_i \equiv \frac{w_i}{\sum_{i=1}^{N} w_i / N}. \tag{3}$$

Next, we consider adding edges by looking at each pair of countries. We randomly generate undirected, unweighted edges between the pair, where the probability of an edge existing between each pair is computed via Eq. 1 [6].

To create realistic weights for these generated edges, we use a heuristic from the distribution of relative edge weights in the cleaned 1996-2020 trade data. From this data, we found that the distribution of the fraction of the smaller of the two country's GDPs very closely follows a reversed log-gamma distribution, as can be seen in Figure 1.

Thus, we generate random edge weights, e_{ij}, according to the following formula:

$$e_{ij} = 10^{-F} \cdot \min(w_i, w_j), \tag{4}$$

where $F \sim \Gamma(6.5571, 0.5794)$. This method of generating random edge weights has the benefit of always being a positive fraction of the smaller of the two countries GDP's, and avoids unrealistic scenarios where one country's trade is several orders of magnitude larger than its nominal GDP.

The creation of network edges and associated weights in the initial year completes the initialization step of the model. The weighted adjacency matrix that is derived from this network is the model's output for the initial year.

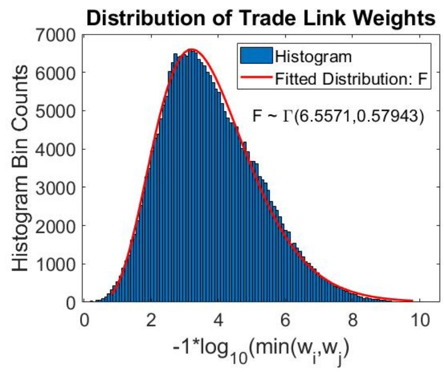

Fig. 1: Reversed distribution of relative trade edge weights in 1996-2020 data, overlaid with fitted Gamma distribution. Analysis and figure creation performed in MATLAB.

3.3 Synthetic Network Growth: Subsequent Years

We generate each subsequent year in the model's output from the previous year's output along with updated GDP data for each country. In a single-pass through the current adjacency matrix, we check the trade relationship status for each pair of countries, with action then taken dependent on this status as follows. If the two countries do not currently have a trade relationship, then an edge is created between them with probability:

$$P_L[x_i, x_j] = \frac{\alpha x_i x_j}{1 + \beta x_i x_j}, \tag{5}$$

where x_i, x_j, and P_L have the meaning of the initialization step's variables, albeit with current year GDP values. The free parameters α and β are allowed to differ from the initialization parameters α_0 and β_0 as discussed in Section 4. Weights are likewise assigned to any newly created edges using Eq. 4 with the current year's GDP values.

If a trade relationship (i.e. weighted edge) already exists for the country pair under inspection, the edge is subjected to random deletion. The probability of edge deletion, P_D, is a function of the current edge weight and the two countries' respective GDPs:

$$y \equiv \log_{10}\left(\frac{e_{ij}}{w_i + w_j}\right), \tag{6}$$

$$P_D[x_i, x_j] = \begin{cases} 10^{(ay^2+by+c)}, & \text{if } y \geq 10^{-10} \\ 0.36, & \text{otherwise.} \end{cases} \tag{7}$$

We heuristically determine the parameters $a = -0.0483$, $b = -0.961$, $c = -5.223$, and the general shape of Eq. 7 from the 1996-2020 trade data. First, we compute overlapping histograms of the different years showing the count of trade relationships that are maintained versus trade relationships which are terminated in a year-over-year fashion, stratified by the value of y from Eq. 6. We then plot the ratio of bin counts as a function of y, and determine that these values follow a roughly quadratic curve in the log-space of these ratios.

Fitting a curve to the logarithm of the 10^{th} through 30^{th} bins yields the heuristic for $y > 10^{-10}$ in Eq. 7, summarized graphically in Figure 2. Note that only the 10^{th} through 30^{th} bins were used since the other bins had too few instances of trade continuing or stopping for the ratio to be statistically meaningful. These 10^{th} through 30^{th} bins roughly correspond with the range of orange bins which are visible in the top portion of Figure 2. For values of $y < 10^{-10}$, we simply set the the probability equal to the maximum of the fitted curve, which is approximately 0.36 in the second plot of Figure 2.

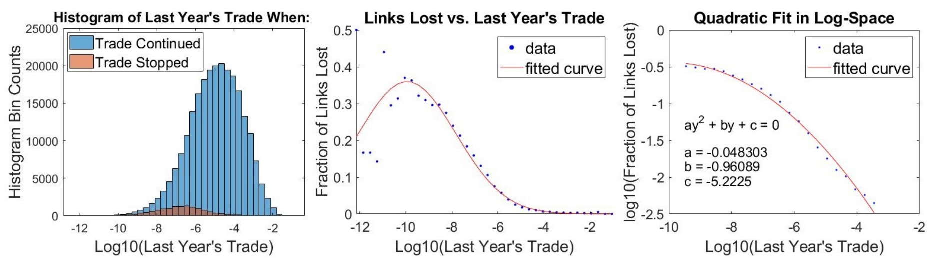

Fig. 2: Left: Histogram of previous year's relative trade value when trade continued and when trade stopped. Middle: Fraction of edges lost taken by computing ratio of histogram bin counts from left image. Right: Log-space of fraction of edges lost for the 10th through 30th bins in the histogram at left, along with fitted quadratic curve. Figures and quadratic fit were generated using MATLAB.

If a trade relationship already exists for the country pair under inspection, and it is not identified for deletion, the weight is slightly adjusted to account for year-over-year changes in the GDPs of the associated countries. We define the relative change in GDP between countries i and j, r_{ij}, as:

$$r_{ij} = \frac{(\min(w_i(y), w_j(y)) - \min(w_i(y-1), w_j(y-1))}{\min(w_i(y-1), w_j(y-1))}, \tag{8}$$

where $w_i(y)$ is the GDP of country i in the current year y, and $w_i(y-1)$ is the GDP of country i in the previous year. If $r_{ij} > 5\%$, then the current edge weight, e_{ij}, is increased by a factor uniformly drawn from the range $[1, 1+2r_{ij}]$. If $r_{ij} < -5\%$, then e_{ij} is decreased by a factor uniformly drawn from the range $[1+2r_{ij}, 1]$. Otherwise, the change in GDP is relatively minor, and so the weight is adjusted by a factor uniformly drawn from the range $[0.95, 1.05]$. This effect is to increase trade weights when GDP growth allows, reduce trade weights when GDP shrinks, and randomly jitter trade weights when GDP remains relatively constant.

3.4 Algorithmic Performance

The synthetic model takes as input a matrix of GDP data of size $N \times M$, where each of the N rows represent individual countries, and each of the M columns represent years.

The output of the synthetic model produces an $N \times N \times M$ data cube, where each $N \times N$ slice along the third dimension represents the weighted, symmetrical adjacency matrix for the WTW for a given year. In this sense, the output is the time-varying adjacency matrix of the graph representing the synthetic WTW.

The initialization and iterative loop are constructed such that each entry in the output matrix is computed exactly once. Since the computation of each entry is deterministic, the worst-case computation time for each entry of the output is a bounded constant. The overall algorithm thus possesses a worst-case time complexity of $O(N^2M)$. Since the initialization and iterative loop computations only require the current and previous year's state information, intermediate variables used in the algorithm have a space complexity bounded by the smaller of $O(N^2)$ and $O(NM)$. Since each of these bounds is smaller than the size of the $N \times N \times M$ output, the overall space complexity of the algorithm is also $O(N^2M)$.

As a reference point for future users, we ran the algorithm using our reference data set ($N = 162$, $M = 25$) 100 times on a Windows 10 personal computer with an i5-4670 (3.40 GHz) CPU and 32 GB of RAM. Chosen model parameters were $\alpha_0 = 220$, $\beta_0 = 80$, $\alpha = 200$, $\beta = 80$. The mean execution time for the algorithm with these inputs was found to be 0.519 seconds, with a standard deviation of 0.00907 seconds.

3.5 Connectivity Enforcement

The MATLAB implementation of the model was programmed supporting an option that requires all nodes to be connected during each time-step of the output. When selecting this option at the end of each time-step, the model merges two different components by randomly creating a bridge between two nodes in each component. The weight for the bridge is computed using Eq. 4. This merging processes is repeated until only a single connected component remains.

Given the iterative nature of this process, and the relatively expensive computation of the connected components during each iteration, this option has the potential to increase the run-time of the algorithm. However, we found that using parameter values of α_0, α, β_0, and β such that the synthetic WTW degree distribution approximates the real-world WTW distribution, there was no appreciable difference in actual execution time. This is due to the rarity of randomly obtaining multiple connected components when using the parameter values that approximate the real-world WTW.

4 Results and Analysis

We organize the analysis of our model into four parts: initial construction, GDP parameters, simulation parameters, and topological structure analysis.

4.1 Initial Construction

Our first attempt at creating the initialization of the model from the 1996 data yielded a smaller network characterized by a depressed degree distribution, especially among the high-degree nodes. This result is expected given that the edge construction probabilities at each time step, stemming from chosen constants α and β, are predicated on an existing representative network to add and remove edges. In the case of the base year however, the synthetic network is only constructed by adding edges to an empty graph of the appropriate order. Without data prior to the initial year, there are no existing edges to maintain, and so trade-link creation probabilities during initialization must be higher to compensate. To achieve this, we establish alternate values for α and β for model initialization. We refer to these parameters as α_0 and β_0.

Given the results from [6], we use the 1996 GDP distribution as the fitness variable to specify the apparent topological characteristics of the real network. We then obtain optimal values for α_0 and β_0 by minimizing the sum of the degree distribution error through the Nelder-Mead simplex algorithm [10]. While any network metrics might be optimized, we observe that reproducing the approximate degree distribution tends to align the other topological characteristics of the synthetic network with the real network. The values that optimize the degree distribution fit are $\alpha_0 = 220$ and $\beta_0 = 80$.

4.2 GDP Parameters

We generate realistic GDP data using parameters measured from the real-world GDP data. Using a log-normal distribution model for GDP data, we computed the mean and standard deviation of this log-normal distribution (in log terms) for each year of available data. The results are presented in Figure 3. As shown, the base year has a mean of approximately 23.2 (approximately \$11.9B in nominal terms), and standard deviation of 2.46. The growth in the mean is approximately linear over the 25 year data set, and the standard deviation remains within a stable range between 2.35 and 2.47. Thus, we generate the base year's synthetic GDP data by randomly selecting values from the 1996 log-normal distribution for each country. Each subsequent year's GDP is generated by applying a random growth factor to the previous years' GDP value.

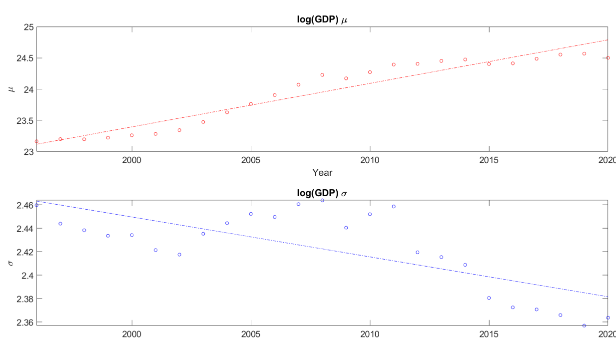

Fig. 3: Top: Logarithm of mean GDP by year between 1996 and 2020, with fitted linear curve. Bottom: Logarithm of standard deviation of GDP by year between 1996 and 2020, with fitted linear curve.

4.3 Simulation Parameters

Due to the randomness inherent in our method with respect to GDP parameters and edge weights, we simulated the synthetic WTW multiple times to avoid bias caused by statistical outliers. Using experimentally determined parameter values of $\alpha_0 = 200$ and $\beta_0 = 80$, we computed a running update of the mean and standard deviation of the parameters in Table 2 after each iteration. We found no significant changes occurred in the Table 2 entries after 30 iterations, and so the simulation was halted at that point.

4.4 Topological Structure Analysis

Tables 1 and 2 show statistics by year for the WTW network using real data and synthetic networks, respectively. We apply the initial construction parameters to 1996 and generate subsequent years iteratively via our methodology for adding and removing trade links. We average the synthetic network statistics of 30 simulations. Each row presents the statistics for each time slice of the temporal network, and global statistics are provided at the bottom of each table. As expected, the year-to-year comparison provides more meaningful data than the average data, where each network in our model is a time-dependent network stimulated by GDP growth and random processes.

The defining feature of the synthetic network is the annual increase in the number of edges that both approximates the actual growth in world trade links and preserves the remaining network topology. Many network characteristics are well-preserved, such as: average degree, average shortest path, average clustering coefficient, and the maximum k-core, with standard deviations as appropriate. We observe from Tables 1 and 2 that the synthetic network presents excellent approximations in each of the relevant statistics.

Figure 4 shows the real and synthetic degree distributions for each year.

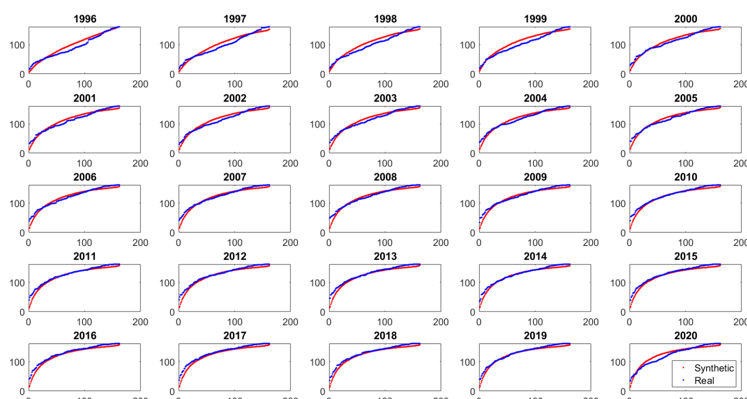

Fig. 4: Degree distribution of the real WTW (in blue) of the retained 162 countries and the synthetic WTW (in red) created from the bootstrapped 1996 GDP values.

We note that the synthetic data aligns well with the real data given the long time frame of the model. Note that the recreation of the degree distribution is not the sole indicator of a successful synthetic network. As a basis for comparison, we created a random MR-configuration graph for each year's actual degree distribution, which preserved none of the other desired topological characteristics. For example, an MR Configuration random network of the 1996 degree distribution produced an average shortest path length of 1.59, an average clustering coefficient of 0.5 and a max K-core of only 44. We observed similarly disparate statistics for the remaining years.

The tabulated statistics are derived from unweighted, undirected versions of the network. Due to the extremely high density of the network and the extremely high probability of reciprocal trade links, the topological characteristics of the weighted, directed networks persist in their simplified reductions [12]. However, we inspected several aspects of the weighted networks for verification and consistency. For example, in 1996, the total value of traded goods was \$999.3B out of a possible "worldwide" (the 162 countries in the data set) GDP valuation of \$30.6T, while an average of 30 trials in the synthetic network for that year estimated \$885.7B of trade out of a simulated \$42.0T. Subsequent years produced similar metrics well within reasonable parameters for our randomized modelling of GDP weights.

Table 1: Network statistics of the real WTW from 1996 to 2020.

Year	N	E	Density	μ Degree	σ Degree	μ Shortest Path	σ Shortest Path	μ Clustering Coeff	σ Clustering Coeff	Max k-Core
1996	162	7486	0.287	92.420	39.971	1.426	0.505	0.819	0.143	68
1997	162	7931	0.304	97.914	38.396	1.392	0.499	0.820	0.119	72
1998	162	8195	0.314	101.173	38.052	1.372	0.494	0.829	0.112	77
1999	162	8369	0.321	103.321	37.647	1.358	0.490	0.833	0.102	78
2000	162	9030	0.346	111.481	34.996	1.308	0.471	0.843	0.095	84
2001	162	9102	0.349	112.370	34.702	1.302	0.469	0.844	0.092	85
2002	162	9194	0.353	113.506	34.553	1.295	0.466	0.847	0.086	85
2003	162	9373	0.359	115.716	33.314	1.281	0.459	0.848	0.084	88
2004	162	9540	0.366	117.778	32.599	1.268	0.453	0.853	0.084	89
2005	162	9627	0.369	118.852	32.566	1.262	0.449	0.857	0.080	91
2006	162	9758	0.374	120.469	32.106	1.252	0.077	0.861	0.078	93
2007	162	9960	0.382	122.963	31.069	1.236	0.434	0.868	0.077	96
2008	162	9943	0.381	122.753	31.008	1.238	0.435	0.866	0.075	96
2009	162	10036	0.385	123.901	30.396	1.230	0.431	0.868	0.075	95
2010	162	10220	0.392	126.173	29.626	1.216	0.421	0.875	0.074	99
2011	162	10281	0.394	126.926	29.347	1.212	0.418	0.878	0.073	100
2012	162	10316	0.396	127.358	29.275	1.209	0.416	0.879	0.074	101
2013	162	10405	0.399	128.457	28.790	1.202	0.411	0.882	0.074	102
2014	162	10289	0.394	127.025	29.823	1.211	0.418	0.881	0.072	101
2015	162	10456	0.401	129.086	29.222	1.198	0.408	0.888	0.070	103
2016	162	10518	0.403	129.852	28.726	1.193	0.405	0.890	0.071	106
2017	162	10569	0.405	130.481	28.274	1.190	0.402	0.890	0.070	104
2018	162	10504	0.403	129.679	28.930	1.195	0.406	0.890	0.067	103
2019	162	10399	0.399	128.383	30.187	1.203	0.412	0.889	0.068	102
2020	162	9890	0.379	122.099	33.535	1.242	0.438	0.876	0.065	96
min		7486	0.287	92.420	28.274	1.190	0.402	0.819	0.065	68
max		10569	0.405	130.481	39.971	1.426	0.505	0.890	0.143	106
avg		9656	0.370	119.205	32.285	1.260	0.442	0.863	0.083	92.6
std		859.3	0.033	10.609	3.365	0.066	0.031	0.022	0.0181668	10.4

Table 2: Network statistics for our synthetic WTW computed over 30 iterations.

Year	N	E	Density	μ Degree	σ Degree	μ Shortest Path	σ Shortest Path	μ Clustering Coeff	σ Clustering Coeff	Max k-Core
1996	162	7471	0.298	92.231	42.515	1.422	0.494	0.871	0.141	81
1997	162	7773	0.310	95.968	39.276	1.401	0.492	0.839	0.118	82
1998	162	8196	0.325	101.187	38.587	1.369	0.484	0.848	0.109	87
1999	162	8470	0.338	104.570	37.911	1.349	0.477	0.853	0.103	89
2000	162	8670	0.348	107.040	37.364	1.332	0.472	0.857	0.098	91
2001	162	8832	0.356	109.034	36.913	1.320	0.467	0.861	0.094	93
2002	162	8952	0.362	110.521	36.529	1.311	0.463	0.864	0.092	94
2003	162	9053	0.366	111.768	36.196	1.305	0.460	0.866	0.088	95
2004	162	9151	0.370	112.970	35.919	1.297	0.457	0.869	0.086	96
2005	162	9226	0.373	113.906	35.670	1.290	0.454	0.871	0.085	97
2006	162	9297	0.374	114.773	35.437	1.285	0.452	0.873	0.083	98
2007	162	9358	0.376	115.528	35.233	1.281	0.449	0.874	0.082	99
2008	162	9412	0.378	116.200	35.042	1.276	0.447	0.876	0.080	99
2009	162	9466	0.380	116.861	34.841	1.273	0.445	0.877	0.079	100
2010	162	9506	0.381	117.361	34.679	1.270	0.444	0.878	0.078	100
2011	162	9544	0.384	117.831	34.538	1.267	0.442	0.879	0.077	101
2012	162	9577	0.384	118.232	34.408	1.264	0.441	0.880	0.076	101
2013	162	9611	0.386	118.651	34.278	1.261	0.439	0.881	0.074	102
2014	162	9638	0.386	118.987	34.156	1.259	0.438	0.882	0.074	102
2015	162	9663	0.389	119.298	34.087	1.257	0.437	0.882	0.074	102
2016	162	9688	0.389	119.608	33.980	1.254	0.436	0.883	0.073	102
2017	162	9710	0.389	119.873	33.885	1.253	0.435	0.884	0.072	103
2018	162	9726	0.388	120.075	33.797	1.252	0.435	0.884	0.073	103
2019	162	9748	0.389	120.343	33.734	1.251	0.434	0.885	0.071	103
2020	162	9765	0.390	120.555	33.635	1.249	0.433	0.885	0.071	103
min		7471	0.298	92.231	33.635	1.249	0.433	0.839	0.071	81
max		9765	0.390	120.555	42.515	1.422	0.494	0.885	0.141	103
avg		9180.12	0.368	113.335	35.704	1.294	0.453	0.872	0.086	96.9
std		617.5	0.025	7.623	2.080	0.047	0.018	0.012	0.017	6.4

5 Conclusions and Further Directions

In this work, we presented a GDP growth-based, time-dependent world trade model to forecast future synthetic networks of the WTW. Despite the underlying complexity and geopolitical considerations likely driving real-world behavior, we find that the single macroscopic measure of GDP can adequately drive the time-evolution of the WTW. The model approximates, to a surprisingly high degree of accuracy, the real network degree distribution, clustering behavior, and other relevant topological characteristics.

Future work may find value in more appropriate GDP approximations and edge weight adjustments. While individual countries may slightly differ, global GDP growth rates appear to follow a Cauchy distribution, vice our simplified linear model [14]. This application could improve our model's ability to match real-world data years in the future, as a small error in our approximation could compound year-by-year. Additional consideration of the edge weight adjustments could likewise be fruitful: instead of heuristically tying edge weights to changes in an individual country's GDP, a distributional analysis of edge weights may improve model performance.

It is our hope that this lightweight, intuitive model can be used in the future to demonstrate, reproduce, and analyze the economic effects of major world events to yield further insights into the nature of our complex global trade system.

Bibliography

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. Reviews of modern physics, 74(1):47, 2002.

[2] Guido Caldarelli, Andrea Capocci, Paolo De Los Rios, and Miguel A Munoz. Scale-free networks from varying vertex intrinsic fitness. Physical review letters, 89(25):258702, 2002.

[3] DESA/UNSD, United Nations Comtrade database. Annual goods trade: 1996-2020. Accessed February 3, 2022, https://comtrade.un.org/data/.

[4] Matthew L Elliot, Sanjeev Goyal, and Alexander Teytelboym. Networks and economic policy. Oxford Review of Economic Policy, 35:565–585, Winter 2019.

[5] Alan Frieze and Michał Karoński. Introduction to random graphs. Cambridge University Press, 2016.

[6] D. Garlaschelli, T. Di Matteo, T. Aste, G. Caldarelli, and M. I. Loffredo. Interplay between topology and dynamics in the World Trade Web. The European Physical Journal B, 57(2):159–164, May 2007. arXiv: physics/0701030.

[7] Diego Garlaschelli and Maria I Loffredo. Patterns of link reciprocity in directed networks. Physical review letters, 93(26):268701, 2004.

[8] Diego Garlaschelli and Maria I Loffredo. Structure and evolution of the world trade network. Physica A: Statistical Mechanics and its Applications, 355(1):138–144, 2005.

[9] Petter Holme and Jari Saramäki. Temporal networks. Physics reports, 519(3):97–125, 2012.

[10] Jeffrey C. Lagarias, James A. Reeds, Margaret H. Wright, and Paul E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. SIAM Journal on Optimization, 9(1):112–147, 1998.

[11] A.S.d. Mata. Complex networks: a mini-review. Brazilian Journal of Physics, pages 658–672, July 2020.

[12] Ma Ángeles Serrano and Marián Boguñá. Topology of the world trade web. Phys. Rev. E, 68:015101, Jul 2003.

[13] The World Bank. Gdp (current us$) - world. Accessed February 23, 2022, https://databank.worldbank.org/home.aspx.

[14] Michael A Williams, Grace Baek, Yiyang Li, Leslie Y Park, and Wei Zhao. Global evidence on the distribution of gdp growth rates. Physica A: Statistical Mechanics and Its Applications, 468:750–758, 2017.