Predicting Crisis in the Global Trade Network

Christina Kao, Lili Yang, and Ye Yuan

December 7, 2015

1 Introduction

The web of international trade relationships is a natural domain in which to apply the discipline of network analysis. Network analysis can reveal insights into the structures of trade relationship and interdependencies between trade partners that are not immediately evident in a straight out statistical analysis of trade data. Unsurprisingly, there is a large body of work from the last decade on this very subject. However, most of this previous work has been focused on analyzing the structure of the network and its evolution through time on a global scale. Relative little work has been to analyze how the trade network might be shaped locally by transformative local events - such as natural disasters, violent conflicts, and financial crisis. In particular, an intriguing question arises: Can we build a model that can identify the occurrence of a local crisis simply by examining the trade network?

This is the question that we seek to answer. In the following sections, we summarize and discuss some of the previous literature in the study of global trade networks, as well as papers on statistical and network analysis methodology applicable to such a study. Next, we detail some of our own analysis of the global trade network between the years 2005 and 2014, both on a global and a local level. Finally, we describe our machine learned model for predicting local points of crisis using data and statistic gleaned from the global trade web, and discuss our findings and their implications.

2 Relevant Works

2.1 Weighted Network Analysis and Modeling

In [2], the authors show that the International Trade Network obeys scale-invariance and universality properties through the following findings: (i) the scaled link annual weight distribution of the trade network approximates a log-normal distribution, and remains unchanged from 1948 to 2000 (ii) both the number of nodes and links grew systematically over the 53 years, yet the link density remains roughly constant (iii) power-law growth of the trade strength with gross domestic product is observed as a universal feature, and the exponent is similar for all countries (iv) the number of countries controlling half of the world's trade is shrinking. We adopt these four major findings as the baseline characteristic assumptions for our network, and report our findings in the Analysis section.

2.2 The Evolution of the World Trade Web

In [3], Fagiolo, Reyes, and Schiavo builds on the work of Bhattacharya, Kunal, et al in using weighted network analysis to glean insights from the World Trade Web (WTW) and its evolution through time. Their main conclusions are that, in general, the countries in the WTW tend to have weak trade link, but there exists a small set of countries with very strong links, suggesting that the network can be modeled as a core-peripheral structure. In addition, the authors conclude that these network properties are remarkably stable across time, and robust under different edge weighting methodologies.

They computed a wide range of network properties on the WTW, and the ones we adopted include node degree, node strength, average nearest neighbor degree, cluster coefficient and random walk betweenness centrality. Our results show that our network is consistent with their findings as presented in the Analysis section.

2.3 Superfamilies of evolved and designed networks

[1] presents a new approach to systematically study similarities in the local structure of networks. Through this method, we can recognize the similarity by calculating the triad significance profile (TSP) of small sub-graphs in the network compared to randomized networks. The TSP of a given network is computed by counting the number of triads in the network of each triad type and comparing with counts that would occur in a randomized network with similar degree and density properties. And we applied this method to our network; the results detailed in the Analysis section.

3 Data Acquisition

We acquired our data from http://comtrade.un.org/, an online repository of official global trade statistics maintained by the United Nations, using the site's open API. The UN Comtrade database contains annual trade data from 1962, and monthly trade data from January 2010. We will primarily be focusing our analysis on the most recent decade of data.

The raw data takes roughly the following form: the reporter country, the partner country, the type of trade (import or export), and the trade volume in US dollars. The data is also broken down into product categories under the Harmonized System (HS), an internationally standardized system for classifying traded products, though our initial analysis focused on the total trade volumes across product categories. Finally, the trade network participants can vary from year to year, though we have, on average, data for around 240 countries in any given time frame.

After we acquired the data for a given time frame - year or month, we parsed it into the form of a directed network where the nodes represent the countries. A directed edge from country A to country B represents the flow of goods from country A to country B. The weight of such an edge is computed by summing the value of the exports from country A to country B as reported by country A and the value of the imports received by country B from country A as reported by country B. These volumes are non-inclusive, which is why we compute the sum instead of the average. (The dataset also reports on re-imports and re-exports, but these volumes are relatively insignificant and thus we have opted to omit them from our graph.)

4 Analysis

4.1 Baseline Network Statistics

We used the results from [2] and [3] as baseline assumptions for our network. Through computing and comparing the variables used in the two papers, we discovered that our network possesses the following properties: (1) the link annual weight distribution approximates a log-normal distribution (2) the number of nodes remains roughly the same, yet the number of links resembles a bell curve from 2005 to 2014 (3) resources are becoming more concentrated, as the number of countries controlling half of the world's trade is shrinking compared to 1948 and 2000 (4) the node strength and random walk betweenness centrality distributions of our network show that it can be modeled as a core-peripheral structure.

4.1.1 Weighted Network Analysis and Modeling

In this section, we present our network statistic results with respect to each of the four assumptions in [2]:

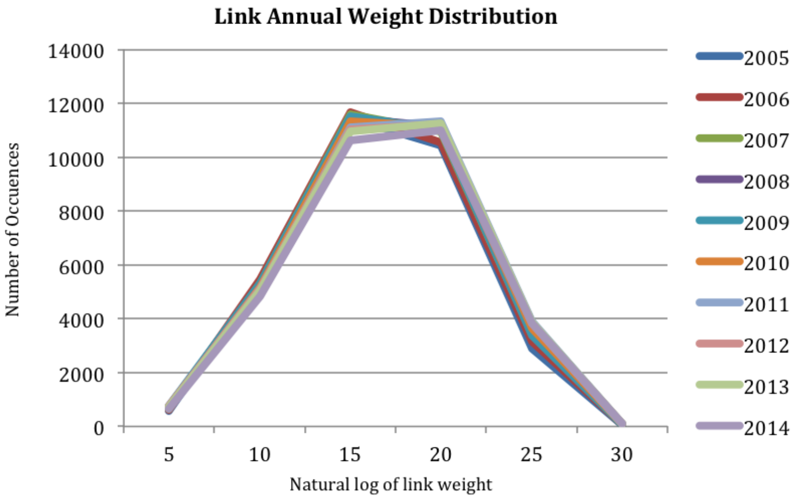
(i) The scaled link annual weight distribution of the trade network approximates a log-normal distribution.

We computed the natural log of the annual link weights, and found that the distribution for our network also approximates a log-normal distribution across all years as shown in Figure 1.

(ii) Both the number of nodes and links grows systematically, yet the link density remains roughly constant.

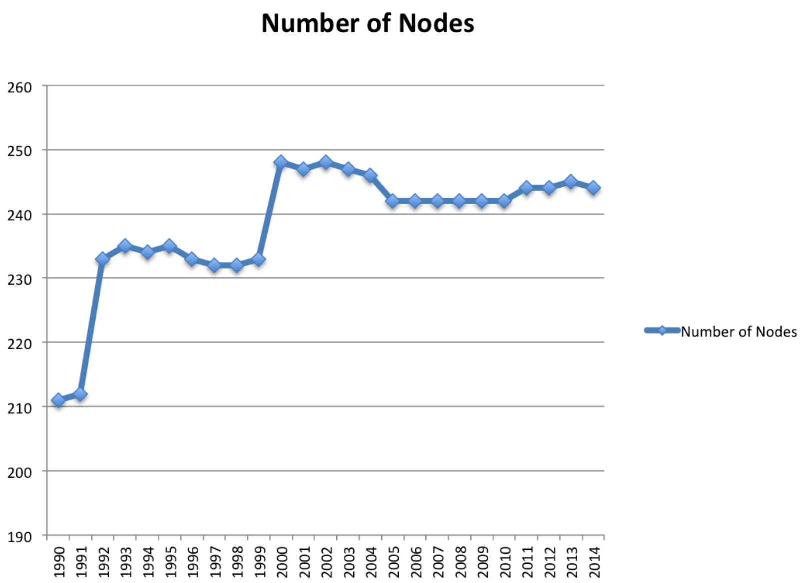
Both the number of nodes and links grew systematically from 1990 - 2000, yet remains roughly constant then on-wards. The lack of growth in the number of nodes is most likely due to the saturation of the trade network - by this point most countries in the world are already part of the global trade network (Figure 2). As for links, there's a slight decrease starting from 2008, most likely due to the global financial and Europe debt crisis (Figure 3). And for link density, it is calculated as [L/[N(N-1)]/2] in [2]. The density remains roughly constant in our network, yet the average value, being around 1.08,

Figure 1: Link annual weight distribution for years 2005 - 2014



is nearly double that of the reference value, 0.52, from 1948 to 2000 in [2]. During the earlier period, there was high growth in both the number of new countries involved in the trade network, and also trade links. However, since most countries were already part of the network by 2000, one potential reasoning behind the high link density is that further growth resulted mainly from new trade relations established among countries.

Figure 2: Global Trade Network - Number of nodes for years 1990 - 2014



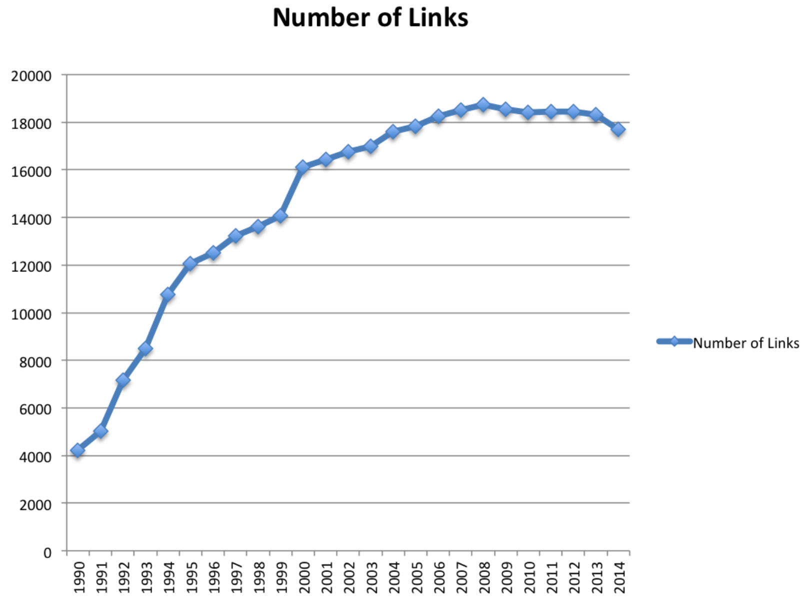
(iii) Power-law growth of the trade strength with gross domestic product as a universal feature.

We computed the the growth of the trade strength for the whole network with gross world product, and the growth of trade strength with gross domestic product for Japan, as we focus on examining trade network changes surrounding the Japan Tohoku earthquake and tsunami in 2011 in the Crisis section. For the whole network, the trend only persisted until 2008, most likely interrupted by the financial crisis. As for Japan, there are positive correlations between the two variables, yet they do not resemble power-law nor possess any clear trends. Japan might not have been the best choice for this analysis due to its economic stagnation over the past 20 years.

(iv) The number of countries controlling half of the world's trade is shrinking.

The number of countries controlling half of the world's resources from 2005 to 2014 remains roughly

Figure 3: Global Trade Network - Number of links for years 1990 - 2014



constant, at around 10. However, compared to 1948 and 2000, the two referenced periods in [2], it has decreased percentage wise, with 1948 at 19%, 2000 at 8%, and 2005 to 2014 at around 5%.

4.1.2 The Evolution of the World Trade Web

Using [3] as a reference, we computed the node degree, node strength, average nearest neighbor degree, cluster coefficient and random walk betweenness centrality of our network, and the results for node strength and random walk betweenness centrality support the conclusion in [3] that the global trade network can be modeled as a core-peripheral structure.

We chose to compute the above five variables, as they are able to indicate the connectivity, assortativity, clustering and centrality properties of networks. The cluster coefficient remains roughly constant across all 10 years, averaging at about 0.85. As for node degree distributions, they are similar across the time frame. The number of nodes gradually increases as the degree increases. And the node strength and random walk betweenness centrality from 2005 to 2014, as shown in Figure 4 and 5, indicate that in general, the countries in the trade network have weak trade links, but there exists a small set of countries with very strong links, suggesting that the network can be modeled as a core-peripheral structure. In addition, these properties are stable across time.

4.2 Triad Significance Profile

We computed the triad significance profile for our network following the procedure described in [1]. Using exactly the same triad categories defined in that paper, we computed the counts of the occurrence of each type of triad in each of the annual trade network between 2005 and 2014. For each triad category i, we computed the z-score by taking: $Z_i = (Nreal_i - < Nrand_i >)/\,\mathrm{std}(Nrand_i)$, where $< Nrand_i >$ and $\mathrm{std}(Nrand_i)$ are the mean and standard deviation respectively of the triad counts in an Erdos-Renyi random graph with the same number of nodes and edges.

We then computed the triad significance profile by normalizing across all of the z-scores, ie $SP_i = Z_i/(\sum Z_i^2)^{1/2}$. You can see our results for the years 2006, 2008, 2010, 2012, and 2014 in Figure 6.

From this, we can make a number of interesting observation. First, notice that there are only two type of triads with positive z-scores - type 6 and type 13. Type 6 corresponds to the subgraph with nodes $N = \{0, 1, 2\}$ and edges $E = \{(0, 1), (1, 0), (0, 2), (2, 0)\}$, ie the triad where there exists mutually directed edges between two of the pairs of nodes. Type 13 corresponds to the subgraph with nodes $N = \{0, 1, 2\}$ and edges $E = \{(0, 1), (1, 0), (0, 2), (2, 0), (1, 2), (2, 1)\}$, ie the triad with all possible directed edges. This suggests that the majority of edges in the trade network are reciprocal. That is, if country A trades

Figure 4: Node Strength distribution for years 2005 - 2014

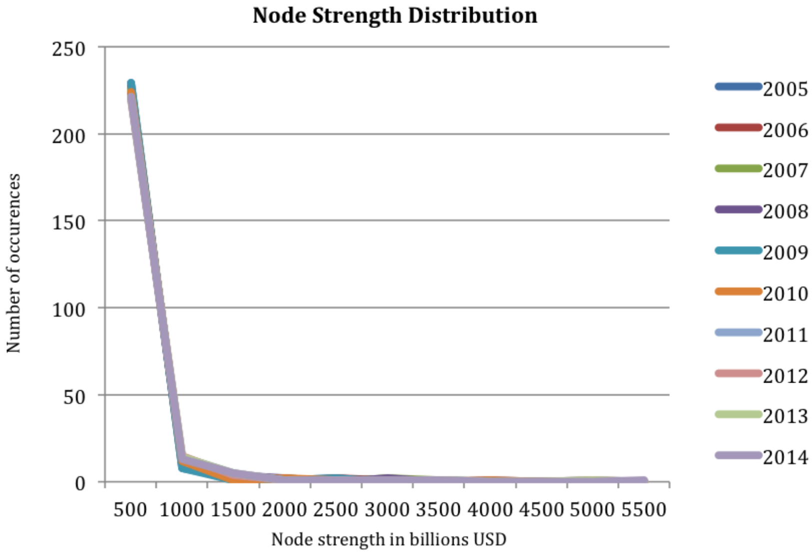
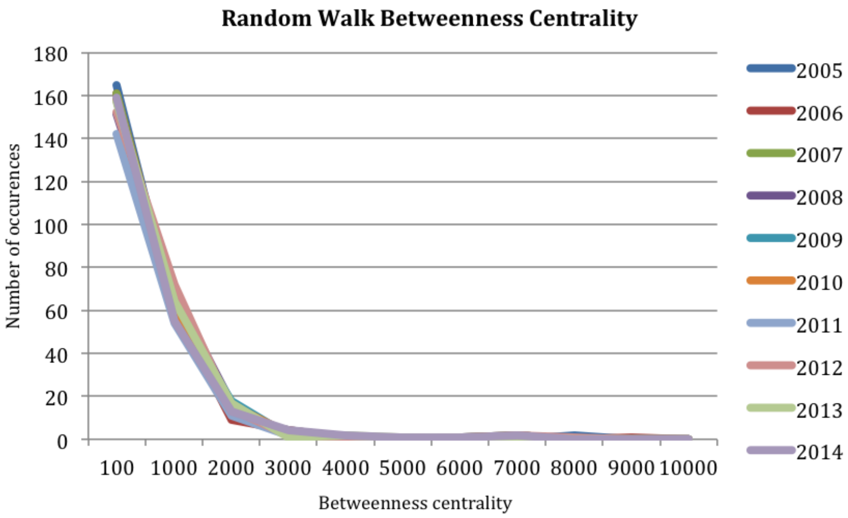


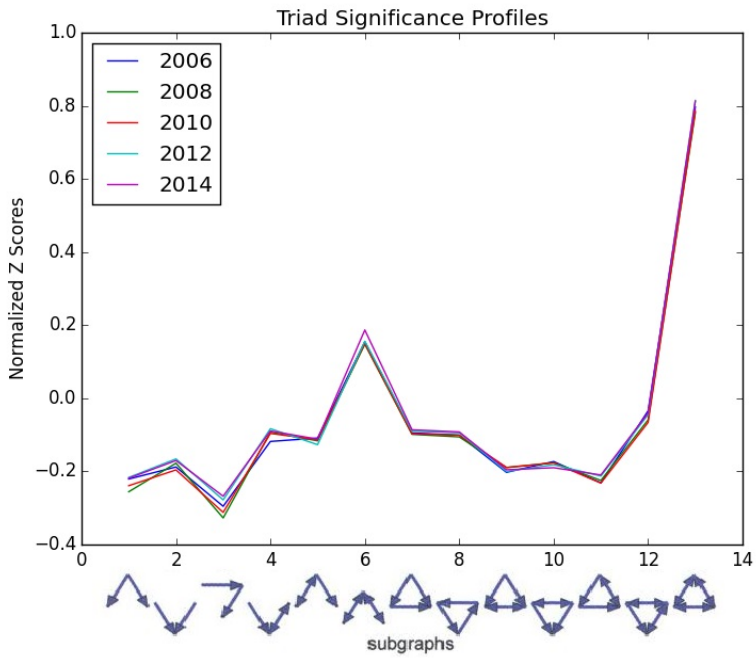
Figure 5: Random Walk Betweenness Centrality distribution for years 2005 - 2014



with country B, then it is like that country B also trades with country A, which tallies with our intuition regarding trade relationships. The prevalence of type 13 triads in particular tallies with what we'd learned in class regarding general signed networks. The friend of a friend is a friend. The trade partner of a trade partner is a trade partner.

Secondly, notice that the triad significance profile for the global trade network doesn't not change significantly from year to year. This is inline with the conclusions of [1], which noted that the triad significance profile, as a technique, can be remarkably robust with respect to changes and fluctuations in the network. It is, however, interesting to note that the largest variance in the significance profiles are for triads of type 1 and 3, which correspond to subgraphs containing only non-reciprocated edges between two pairs of nodes. It is further interesting that the year with the lowest z-scores for those two categories is 2008, which is also the first year of the global financial crisis. This could suggest that non-reciprocal trade relationships of the sort described by type 1 and type 3 are the weakest and most easily broken by crisis. This is something that we intend to explore further as we move forward.

Figure 6: Triad significance profiles for 2006, 2008, 2010, 2012, 2014. Subgraph images taken from [1].



4.3 Crisis

The goal of our project is to try to build a model that can identify the occurrence of a local crisis simply by examining the trade network. As a result, we decided to focus on natural disasters; in particular, earthquakes, floods, and hurricanes and typhoons, the three categories that have sufficient occurrences between 2010 to 2014 (each with approximately 15 events), the period which we have access to monthly trade data to, in order to train the classification model.

4.4 Analysis of a Crisis

We started with the Japan Tohoku earthquake and tsunami that occurred in March of 2011, as it is one of the disasters with the largest impacts. We computed the subgraph for Japan by including all of the country's direct neighbors and links. We then calculated the relevant properties from our baseline network statistics for Japan's sub-network, which include changes in number of nodes and links, link density, node degree, node strength, average nearest neighbor degree and random walk betweenness centrality.

The two properties that showed significant changes after the disaster are node strength and random walk betweenness centrality. There was a 15% decrease in Japan's trade volume in April of 2011, a month after the crisis, as shown in Figure 7 And betweenness centrality was around 3000 to 4000 throughout 2010 and the beginning of 2011, but after the tsunami happened, it dropped through 3000, and in May went down to around 2500 (Figure 8).

5 Predictive Model for Crisis Classification

Using the baseline statistics we've collected and and building upon the analysis that we've done, we've worked on building a machine learned classification model that can successfully identify the occurrences of crisis in a given country at a given time.

5.1 Data

We derived the data used to train our model from our monthly trade data. Specifically, we focused on the 31 countries for which we found the occurrence of a natural disaster between the dates January 2010 and December 2014. Complete lists of countries and disasters are included in the appendix. We

Figure 7: Japan Monthly Node Strength 2010 - 2011

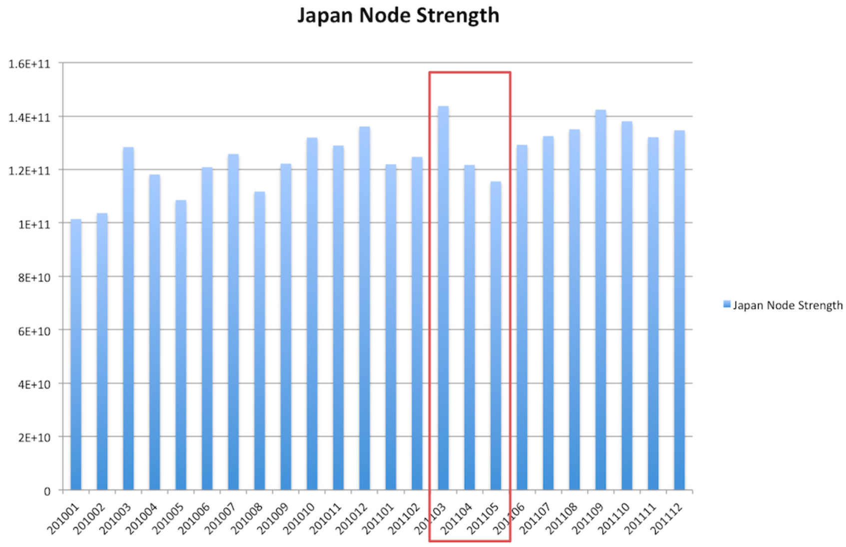
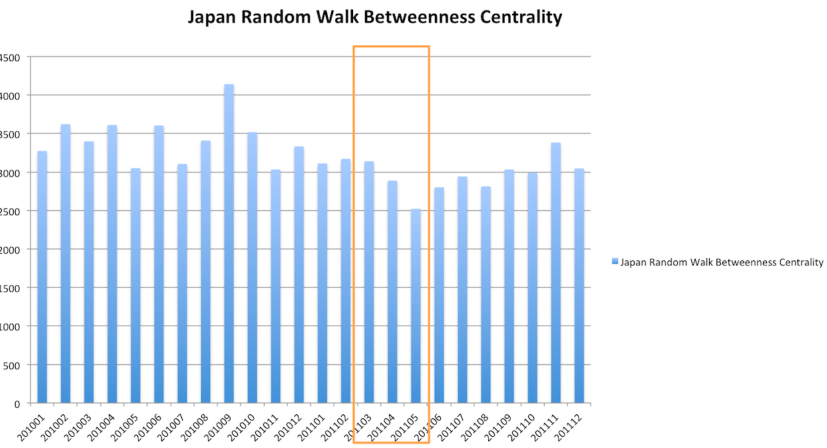


Figure 8: Japan Random Walk Betweenness Centrality 2010 - 2011



considered each data point as a combination of month and country. We labeled the data point as "crisis" if a crisis occurred or was ongoing in that country during that month and "non-crisis" otherwise.

We split our labeled data set into a training set S_{train} (90% of the data) and a test set S_{test} (10%). We assigned data point randomly, though we endeavored to ensure that both our training set and our test set received approximately the same proportion of crisis points. Overall, we have 97 crisis data points and 1391 non-crisis points, ie approximately 6.5% of our data was labeled as "crisis".

5.2 Evaluation

One metric with which we evaluated the success of our model is the overall accuracy - the proportion of data points that were correctly classified. However, since there are considerably more non-crisis points than crisis points, that metric, by itself is insufficient. (In fact, we could achieve nearly a whooping 93.5% accuracy simply by classifying every test point as "non-crisis" but clearly this would not be a good model.)

Consider the goal of our project - it's to correctly identify crisis points. Therefore, more relevant evaluation metrics are the following:

$$precision = \frac{CC}{CC + FP}$$

$$recall = \frac{CC}{CC + FN}$$

Where CC = # of correctly classified crisis points, FP = # of non-crisis points erroneously classified as crisis points (false positives), and FN = # of crisis points erroneously classified as non-crisis points (false negatives). We selected features to maximize the F1 score, computed as the following:

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

5.3 Baseline

The proportion of our data set that's labeled as "crisis" is approximately 0.065. As a base of reference, if we go through our test set and randomly assign "crisis" labels with probability 0.065, we would achieve an overall accuracy of approximately 0.880 and a F1 score of 0.032. Our goal is for our model to improve upon these scores.

5.4 Methodology

To train our model, we used the logistic regression component of the python library scikit. Since our data set is heavily unbalanced, ie there are considerably more non-crisis points than there are non-crisis points, we weighted every crisis point in our training set in order to produce better results. The ratio of non-crisis points to crisis points is approximately 14. As such, we experimented with weight values between 9 and 14.

See appendix 8.3 for a full list of all of the features that we experimented with. For each of the features, we computed the year over year value of the listed statistic. For example, the node degree for Japan on January 2012 is the node degree of Japan on January 2012 minus the node degree of Japan on January 2011. We chose to use the year over year difference (versus, say, the month over month difference) in order to minimize the impact of seasonal effects. We experimented exhaustively with combinations of 1, 2, 3, 4, 5, 6, 7, and 8 features.

5.5 Results

For weight values of 9, 10, 11, 12, 13, and 14, we experimented with all combinations of one, two, three, four, five, size, seven, and eight features. For each combination of weight value and number of features, we picked a best set of features by F1 score. Table 1 shows the F1 scores computed on the test set of these best sets of features. Table 2 shows the overall accuracy on the test set of these same sets of features. As we can see from these tables, we get the best results using a weight value of 11 and the

Table 1: Logistic regression results for best sets of features based on F1 scores

Weights	F1 Scores for Best Sets of Features							
	One	Two	Three	Four	Five	Six	Seven	Eight
9							0.333	0.353
10	0.143	0.24	0.261	0.286	0.286	0.323	0.333	0.316
11	0.188	0.32	0.333	0.348	0.364	0.348	0.32	
12	0.15	0.267	0.276	0.308	0.323	0.333	0.303	
13	0.216	0.244	0.263	0.256	0.263	0.270		
14	0.246	0.214	0.222	0.219				

following set of five features: # of triads of type 1, # of triads of type 4, # of triads of type 5, # of triads of type 9, and random walk betweenness centrality. With these features, we achieve an accuracy of 0.894 and a F1 score of 0.364. Let's call this Model 1. Our second best results come from using a weight value of 9 and the following set of eight features: # of triads of type 3, # of triads of type 4, # of triads of type 5, , # of triads of type 8, # of triads of type 9, link density, node degree, and random walk betweenness centrality. Let's call this model 2. Table 3 shows the coefficients that we computed for model 1 and model 2.

There's a few observations that we can make from these results. First, notice that random walk betweenness centrality shows up in both models with strong coefficients. Earlier on, in our local analysis

Table 2: Overall accuracy for same sets of features

Weights	Overall Accuracy for Best Sets of Features							
	One	Two	Three	Four	Five	Six	Seven	Eight
9							0.909	0.917
10	0.818	0.856	0.871	0.848	0.886	0.841	0.848	0.902
11	0.803	0.871	0.879	0.886	0.894	0.886	0.871	
12	0.742	0.833	0.841	0.864	0.841	0.848	0.826	
13	0.780	0.765	0.788	0.780	0.788	0.795		
14	0.674	0.667	0.682	0.621				

Table 3: Coefficients for Best Models

Features	Model 1	Model 2
# of triads of type 1	-0.10230378	
# of triads of type 3		0.03987966
# of triads of type 4	0.20856831	0.24633773
# of triads of type 5	0.11942822	0.14705058
# of triads of type 8		-0.19283461
# of triads of type 9	-0.02579486	0.07578571
link density		-0.0496448
node degree		-0.2991632
random walk betweenness centrality	-0.18624342	-0.15730376

of the Japan earthquake and tsunami, we noted that random walk betweenness centrality was one of the two properties that appeared to be mostly strongly affected by the disaster. This is consistent with those results as well as our intuition. Other countries are less likely to send their goods through a country that's been affected by disaster. The other property we noted - node strength - does not show up in either model. However, from our Japan graphs we can see that node strength behaved very similarly to random walk betweenness centrality during the crisis period. It is possible that this correlation ensured that only one of the properties was required for our model.

Second, we observed in our analysis of triad profiles that triads of type 1 and type 3 appeared to be the most sensitive to the 2008 financial crisis. Sure enough, we see both types of triads show up as features in our models - type 1 in model 1 and type 3 in model 2.

Somewhat surprisingly, we see triad of type 4 and 5 show up in both models with strong coefficients. However, upon consideration, this is really not such a surprising result. Recall our analysis of triad profiles that the two types of triads that occur most frequently (relative to their occurrence in a random graph) are triads of type 6 and type 13. Notice also from the subgraphs depicts at the bottom of Figure 6 that we would only require a single link to break in a triad of type 6 to get a triad of type 4 or type 5 instead. Is it really so surprising that we would see an increase triads of type 4 and type 5 in the event of a crisis?

5.6 Error Analysis

Through examining the individual entries of correct predictions, false positives and false negatives of our two sets of best performing features, we found that (1) among the disasters, earthquakes are most often correctly identified while the false negatives comprise mainly of floods; and (2) more than half of the false positives correspond to actual disasters, indicating flaws in our manual disaster identification and data set construction process, but strength in our prediction model.

(1) Differences in results by disaster categories

For our 5 feature set, all of the correct crisis month predictions are earthquakes, and 80% of false negatives are floods. And for the 8 feature set, all of the correct predictions are earthquakes, while 66.7% of false negatives are floods. This could indicate that either earthquakes have greater impact, or result in more consistent changes on the trade network, hence the better prediction results.

(2) Prediction model crisis identification

Among the false positives for both feature sets, more than half correspond to actual disasters or civil unrest, and we listed some examples in Figure 9. When constructing our data set, we identified and labeled the crisis months manually, and only looked at earthquakes, floods, hurricanes and typhoons, therefore our data set is not comprehensive of all disasters.

Figure 9: Model Prediction - "False" Positives and Corresponding Events

Country	Date	Label	Predicted Label	Events Found
Nigeria	2012/01	0	1	Several terrorist attacks on churches, governments and businesses in January
Romania	2012/01	0	1	Start of a 3 year long civil unrest and demonstration in Romania
Mexico	2012/02	0	1	Moderate earthquake in Mexico
Peru	2012/10	0	1	Strong earthquake in June and landslide in October
Portugal	2014/01	0	1	Affected by Cyclone Anne in the beginning of January
India	2014/06	0	1	Start of seasonal monsoon rain in India
Nepal	2014/12	0	1	Might have been impacted by landslides and floods that occurred in Aug 2014

6 Summary

Our study shows the baseline statistics of the global trade network between years 2005 - 2014 to be roughly consistent with previous research findings, with a few exceptions, such as growth stall in number of nodes and links, lack of correlation between GDP and trade strength and triad type changes, that most likely resulted from the maturation of the trade network, and also the global financial and Europe debt crisis. And our machine learning model is able to identify local crisis with some success - generating accuracy and F1 scores well above the random model, and selecting features that are roughly consistent with our initial analysis on the triad significance profile and Japan's earthquake and tsunami. More importantly, it identified disasters and other impactful events that we failed to label in our original data set. For future work, having a more exhaustive crisis data set would most likely improve the model's accuracy. We could also consider separating out the data by disaster categories, and combining correlated features, for instance link density and node degree, to aggregate the network change effects.

7 References

References

[1] Milo, Ron et al. "Superfamilies of Evolved and Designed Networks." *American Association for the Advancement of Science* 303.5663 (2004): 1538-542. Web. 13 Jan. 2010.

[2] Bhattacharya, Kunal, et al. "The international trade network: weighted network analysis and modelling." Journal of Statistical Mechanics: Theory and Experiment 2008.02 (2008): P02002.

[3] Fargiolo, G., Reyes, J., and Schiavo, S "The evolution of the world trade web: a weighted-network analysis." Journal of Evolutionary Economics 20 (4), 479-514

[4] Smith, David A., and Douglas R. White. "Structure and dynamics of the global economy: network analysis of international trade 1965–1980." Social forces70.4 (1992): 857-893.

[5] Hertel, Thomas W., and Thomas Warren Hertel. Global trade analysis: modeling and applications. Cambridge university press, 1997.

[6] Mahutga, Matthew C. "The persistence of structural inequality? A network analysis of international trade, 1965–2000." Social Forces 84.4 (2006): 1863-1889.

[7] Kim, Sangmoon, and Eui-Hang Shin. "A longitudinal analysis of globalization and regionalization in international trade: A social network approach." Social Forces 81.2 (2002): 445-468.

[8] Watts, D. J. and Steven Strogatz (June 1998). "Collective dynamics of 'small-world' networks". Nature 393 (6684): 440–442.

8 Appendix

8.1 List of Natural Disasters

Earthquake
Haiti, Jan 12, 2010 (Crisis Month: Jan, Feb 2010)
Chile, Feb 27, Mar 11, 2010 (Crisis Month: Mar, Apr 2010)
Mexico and Southern California, April 4, 2010 (Crisis Month: Apr, May 2010)
Indonesia, Oct 25, 2010 (Crisis Month: Nov, Dec 2010)
Bonin Island, Japan, Nov 30, 2010 (uninhabited, potentially not much impact)
Tohoku Japan (+ Tsunami), Mar 11, 2011 (Crisis Month: Mar, Apr 2011)
New Zealand ChristChurch Earthquake, Feb 11, 2011 (Crisis Month: Feb, Mar 2011)
Turkey Van Earthquake, Oct 23, 2011 (Crisis Month: Nov, Dec 2011)
Iran Earthquake, Aug 11, 2012 (Crisis Month: Aug, Sep 2012)
Afghanistan Earthquake, Jun 11, 2012 (Crisis Month: Jun, Jul 2012)
Philippines Samar Earthquake, Aug 31, 2012 (Crisis Month: Sep, Oct 2012)
Pakistan Earthquake, late Sep, 2012 (Crisis Month: Oct, Nov 2012)
Chile Earthquake, Apr 1, 2014 (Crisis Month: Apr, May 2014)
Floods
Portugal, Feb 22, 2010 (Crisis Month: Mar, Apr 2010)
Tennessee, Apr 30, 2010 (Crisis Month: May, Jun 2010)
Bangladesh, June 15, 2010 (Crisis Month: Jun, Jul 2010)
Pakistan, July 26, 2010 (Crisis Month: Aug, Sep 2010)
Rio de Janeiro, Apr 5, 2010 (Crisis Month: Apr, May 2010)
China, May 10, 2010 (Crisis Month: May, Jun 2010)
Philippines Flood, Late Dec 2010 - mid January 2011 (Crisis Month: Jan, Feb 2011)
Thailand Flood, 25 July 2011 – 16 January 2012 (Crisis Month: Aug 2011 - Feb 2012)
Beijing, China, July 2012 (Crisis Month: Jul, Aug 2012)
Russia, July 2012 (Crisis Month: Jul, Aug 2012)
UK and Ireland, April 2012 - early 2013 (Crisis Month: Apr 2012 - Feb 2013)
Loreto, Peru, Feb - Mar 2012 (Crisis Month: Feb, Mar, Apr 2012)
Nigeria, July - October 2012 (Crisis Month: Jul - Nov 2012)
North Korea, July 19 - Sep 17 2012 (Crisis Month: Jul - Oct 2012)
Manila Flood, Philippines, Aug 1 - 8, 2012 (Crisis Month: Aug, Sep 2012)
Romania, May 16 - Jun 12, 2012 (Crisis Month: May, Jun, Jul, 2012)
Fiji, Jan 21 - Feb 12, 2012 (Crisis Month: Feb, Mar 2012)
Nepal, May 5, 2012 (Crisis Month: May, Jun 2012)
Pakistan, early September 2012 (Crisis Month: Sep, Oct 2012)
Balkan Floods, May 2014, Serbia, Bosnia and Herzegovina and Croatia (Crisis Month: May, Jun 2014)
Hurricanes/Typhoons
Hurricane Tomas, Category 2, Costa Rica / Haiti, Oct. 29 - Nov 7, 2010 (Crisis Month: Nov, Dec 2010)
Hurricane Frank, Category 1, Mexico, Aug 21-28, 2010 (Crisis Month: Sep, Oct 2010)
Typhoon Conson, Category 1, Philippines, July 11-18, 2010 (Crisis Month: Jul, Aug 2010)
Typhoon Megi, Category 5, Philippines, Oct. 12-24, 2010 (Crisis Month: Oct, Nov 2010)
Cyclone Giri, Category 4, Myanmar, Oct. 20-23, 2010 (Crisis Month: Nov, Dec 2010)
Philippines Tropical Storm Washi, Dec 13 - 19, 2011 (Crisis Month: Dec 2011 - Jan 2012)
Hurricane Irene, Eastern US, Aug 21 - 30 2011 (Crisis Month: Sep, Oct 2011)
Hurricane Sandy, Oct 22, 2012 (Crisis Month: Nov, Dec 2012)
Typhoon Haiyan, Philippines, Nov 3, 2013 (Crisis Month: Nov, Dec 2013)
Typhoon Phailin, India, Thailand, Myanmar, Oct 4, 2013 (Crisis Month: Oct, Nov 2013)
Hurricane Ingrid, Mexico, Texas, Sep 12, 2013 (Crisis Month: Sep, Oct 2013)
Philippines Typhoon Hagupit, Nov 30 - Dec 12, 2014 (Crisis Month: Dec 2014)

8.2 List of Countries Examined in Classification Model

Haiti, Chile, Mexico, United States, Indonesia, Japan, New Zealand, Turkey, Iran, Afghanistan, Philippines, Pakistan, Portugal, Bangladesh, Brazil, China, Thailand, Russia, UK, Ireland, Peru, Nigeria, North Korea, Romania, Fiji, Nepal, Serbia, Bosnia, Croatia, Myanmar, India.

8.3 List of Features For Classification Model

For a given country node c and a given date t (month and year), we computed and experimented with the following features. For each feature, we computed year over year values, normalized across the all values for that country.

# of triads of type 1* containing c
# of triads of type 2* containing c
# of triads of type 3* containing c
# of triads of type 4* containing c
# of triads of type 5* containing c
# of triads of type 6* containing c
# of triads of type 7* containing c
# of triads of type 8* containing c
# of triads of type 9* containing c
# of triads of type 10* containing c
# of triads of type 11* containing c
# of triads of type 12* containing c
# of triads of type 13* containing c
Link density in the immediate subgraph** of c
# of links in the immediate subgraph** of c
Degree of node c
Weighted degree/strength of node c
Average nearest neighbor degree of node c
Weighted average nearest neighbor degree of node c
Average nearest neighbor strength of node c
Random walk betweenness centrality of node c

* See the bottom of Figure 6 to mapping of triad type to subgraph image.
** We define the immediate subgraph of c as the subgraph containing c and its immediate neighbors.