

Analysis of Global Core Countries from a Complex Network Perspective

Peijie Zhao*

Faculty of Natural, Mathematical & Engineering Sciences, King's College London, London, UK

*Corresponding author: ericeweilan@163.com

Abstract. This paper analyzes the importance of countries in a globalized world and explores their cooperative tendencies using a complex network approach. By extracting data from The GDELT Project, an undirected weighted global network is constructed, where each node represents a country. The network is found to be scale-free based on the degree distribution. Core countries are identified using different centrality metrics, such as degree centrality, closeness centrality, and PageRank. The occurrence of war is identified as a significant factor influencing the importance of countries like Israel and Ukraine in the network. By considering the types of events, the study examines the cooperative tendencies of countries and finds an overall trend of cooperation rather than conflict among nations, with varying levels of collaboration. Countries involved in conflicts or wars show limited cooperation with other nations during the selected period. The findings suggest that globalization may influence the cooperative tendencies of countries. The paper contributes to the understanding of the importance of countries in a globalized world and their cooperative tendencies using a complex network perspective

Keywords: complex network; globalization; degree distribution.

1. Introduction

Globalization is the process of integrating economics, politics, and cultures, which breaks the borders between nations and transforms the world into a small village [1]. Numerous countries have played significant roles in the globalization process from various aspects. Therefore, exploring these important countries from different angles could enhance our understanding of globalization. Complex networks, as a tool and a unique perspective, have been applied in analyzing important countries, and some studies have demonstrated many useful outcomes [2]. For example, country centrality is considered as an essential indicator of power in research on international relations [3]. This article establishes a complex global network by extracting a segment of global event data to explore core countries in the network and their tendency to cooperate.

The complex network application is extensive and could be utilized in various practical issues. Many researchers have studied various real-world issues through different complex network models. In 2014, A model built a complex public traffic network with three-weights by setting bus lines as network nodes and analyzed the network's balance by considering departure frequency, passenger flow density, and other factors [4]. In addition, another study established a new complex network for evaluating the reputation of virtual organizations in which the nodes can present entities in the Semantic Grid [5]. There is also a work constructed a complex internet financial risk system network by treating finance risk elements as network nodes [6]. The research finds that nodes around the network center are the main objects of financial risk contagion. After some time, in 2022, a study proposed a complex network based on actual COVID-19 data from Portugal, where each node represents one of six regions in Portugal [7]. The authors found that their method could be applied to outbreak management and govern opening or closing borders in a pandemic context. The above literature shows that complex networks can effectively study the influence of crucial nodes on real issues and network structure.

As a result, in this paper, we will apply a complex network to explore the core countries in the world. The rest of the paper is organized as follows. Section 2 will introduce how we process the data

and how a complex network model is designed. In Section 3, some numerical results and discussion are presented. The conclusion is given in Section 4.

2. Methodology

To build a global network, we use the data from The GDELT Project, which monitors the world's news from over 100 languages, and GDELT Event is updated every 15 minutes. Each event in the GDELT Project is recorded in the GDELT Event dataset, which consists of different aspects. More specifically, each entry in the dataset contains the event's time, location, involvement of two actors, the type of event, and other additional information.

2.1. Data Selection

We initially chose the data from 2023/11/23 to 2023/11/29 to build our network. In our global network model, we want to explore the connections between different countries and then explore the core countries. In this paper, we only focus on the data under the ACTOR ATTRIBUTES, EVENT ACTION ATTRIBUTES, and EVENT GEOGRAPHY categories for each entry. In the first step, we will process the data in the following ways:

Firstly, we extract entries from the data that possess different Country Codes. 'CountryCode' in the GDELT Project is stored in a 3-character CAMEO code, so if Actor 1 and Actor 2 have different Country Codes, we can reasonably believe that this event involves two different countries.

Secondly, to avoid duplication for an event, we compare the 'EventCodes' within each data entry, which are CAMEO action codes describing the action Actor 1 performed on Actor 2. If two or more entries have the same 'EventCodes,' we believe these entries record the same event, and we remove duplicate entries and retain only one of those entries. In addition, if two entries have the same Actor1 and Actor2 but have different EventCodes, we interpret them as separate events occurring in the same place.

2.2. Basic Model

Our next step will construct an undirected weighted network based on the remaining data, where the weight represents the number of events between two countries. The nodes will be labeled with Actor1's and Actor2's Country Codes, and Actor1 and Actor2 within one event will be connected. We will also build a matrix showing the number of events between countries.

After constructing the global network, we will first determine the type of this network by observing the degree distribution. After that, to explore the critical country in the network, we will apply different metrics, such as degree centrality, closeness centrality, and page rank method.

2.3. Network Analysis

By observing the degree distribution, we can find the type of our global network. In addition, different metrics can evaluate nodes' importance from various perspectives. We choose degree centrality in this paper because it can directly show the number of events related to one country. Also, closeness centrality could calculate the average distance between nodes, and PageRank considers both the quality and quantity of connections of a node. Both these metrics help assess the influence and importance of nodes in the network.

The degree distribution is one of the most extensively studied properties of a graph or network, which influences the structure of a graph and the processes that occur on it [8]. We can understand how nodes connect to others in the network through degree distribution. In addition, we can observe which nodes are highly connected and which nodes are isolated or have fewer connections.

Degree centrality is the simplest method to measure centrality, defined by the number of connections attached to each node [9]. The degree centrality for node i could be defined as follows,

$$k_i = \sum_{j=1}^N A_{ij} \quad (1)$$

Where N is the total number of nodes in the network, and A_{ij} is the element in the i -th row and j -th column of the adjacency matrix A .

Closeness centrality is defined as the average distance of each node to all other nodes, and it stands for the convenience and ease of connections between the focused node and the other nodes [9, 10]. The closeness centrality of node i could be defined as,

$$x_i = \frac{N}{\sum_{j=1, j \neq i}^N d_{ij}(A)} \quad (2)$$

Where d_{ij} is the shortest path between node i and j .

PageRank operates on the principle that a webpage's importance can be evaluated based on the webpages that link to it [11]. As a result, the importance of a page or node depends on how many links point to it and the importance of the pages or nodes that link to it. A higher PageRank value implies a superior ranking in search engine results, as it is seen as having more authority or influence. In addition, although the PageRank was traditionally applied on directed graphs, it could also be applied to undirected graphs [12].

2.4. Cooperative Tendency

In the previous steps, we used network structure or properties to determine the critical country nodes. However, a country's attitudes towards other nations also play an essential role within the global sphere. Categorizing events into cooperation and conflict could help us to evaluate the cooperative tendencies of these nodes within the network.

In this step, we will use the 'QuadClass' from the data to determine the status of a country, whether it is in a cooperative or conflicting state. We will assign values in 'QuadClass' as follows: '1' represents Verbal Cooperation, assigned a value of +1; '2' represents Material Cooperation, assigned a value of +2; '3' represents Verbal Conflict, assigned a value of -1; and '4' represents Material Conflict, assigned a value of -2.

After processing the data from the selected period, 192 countries are within our network. In the next step, we will create a 192*192 matrix for further study.

3. Results and Discussion

When obtaining the degree of all nodes, we will take the logarithm of each value and arrange them in descending order, trying to fit them using a curve. Fig. 1 shows that the fitted curve demonstrates a linear trend. In addition, based on this observation, we can conclude that our network is a scale-free network. A scale-free network means that some hubs have seemingly unlimited links while others only have a few connections [13]. The outcome of the above graph is not coincidental, as the degree distribution in real networks is more likely to power law distribution.

In scale-free networks, crucial nodes could significantly influence the network. So, in the next step, we will find and analyze important nodes in our network through metrics we introduced before.

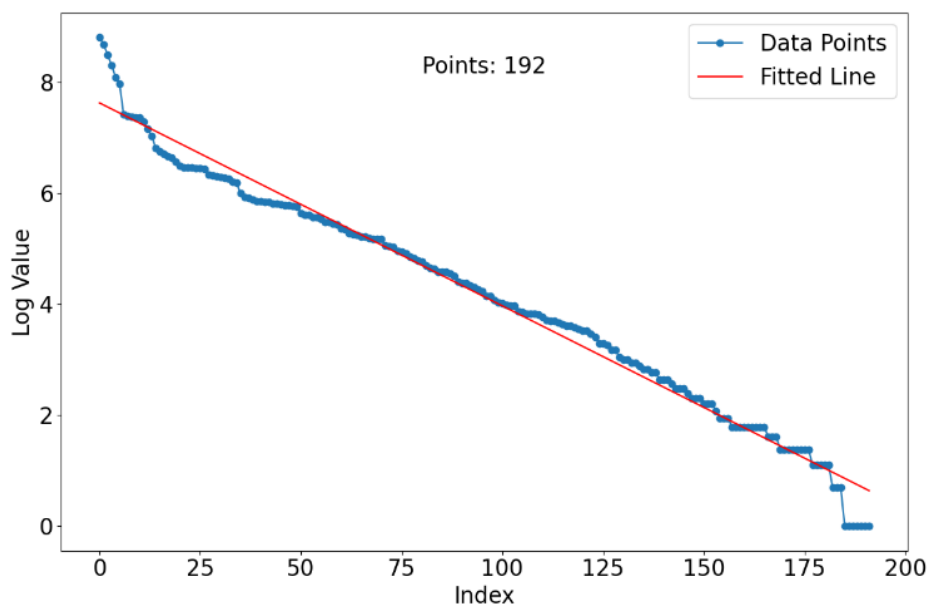


Fig. 1 Descending orders of log values of degrees.

3.1. Importance of the Nodes

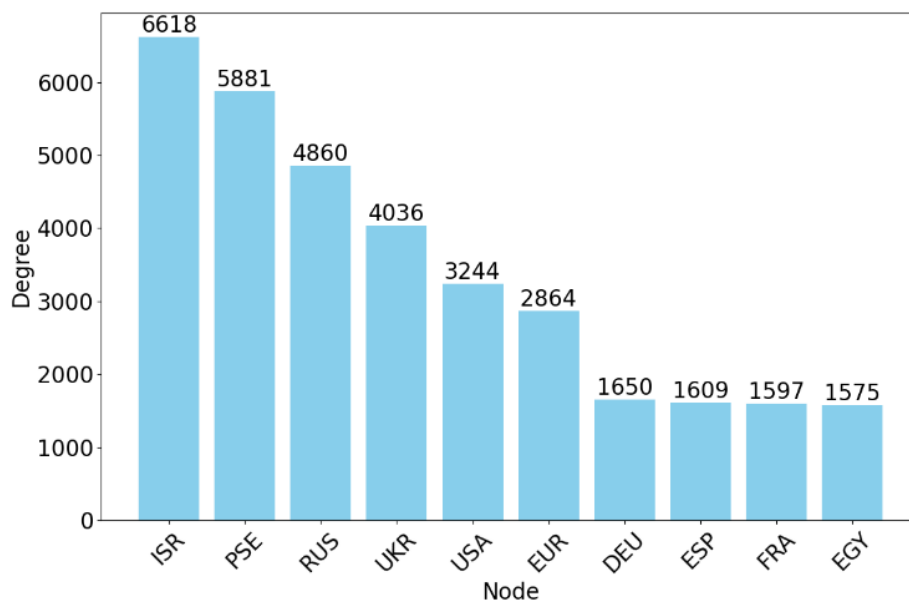


Fig. 2 Top 10 nodes by degree.

From the Fig. 2, we can find that the ISR (Israel) has the largest degree centrality value, and PSE (Palestinian Occupied Territories) has the second highest degree. In addition, RUS (Russia) and UKR (Ukraine) share a similar degree range, ranging between 4000 and 5000. Moreover, there is a notable decrease from the seventh position (DEU), and the degrees are closely aligned in size, around 1600 starting from the seventh.

ISR and PSE have higher degrees, meaning more events involve these two countries. This observation aligns with the fact because, during the chosen period, Israel and Palestine were engaged in the war, resulting in more media or news reporting events about these two countries. Similarly, RUS and UKR were also in a war. However, there were fewer news reports on RUS and UKR due to the extensive time span between the starting time of their conflicts and our selected period. As a result, their degrees are lower than ISR's and PSE's.

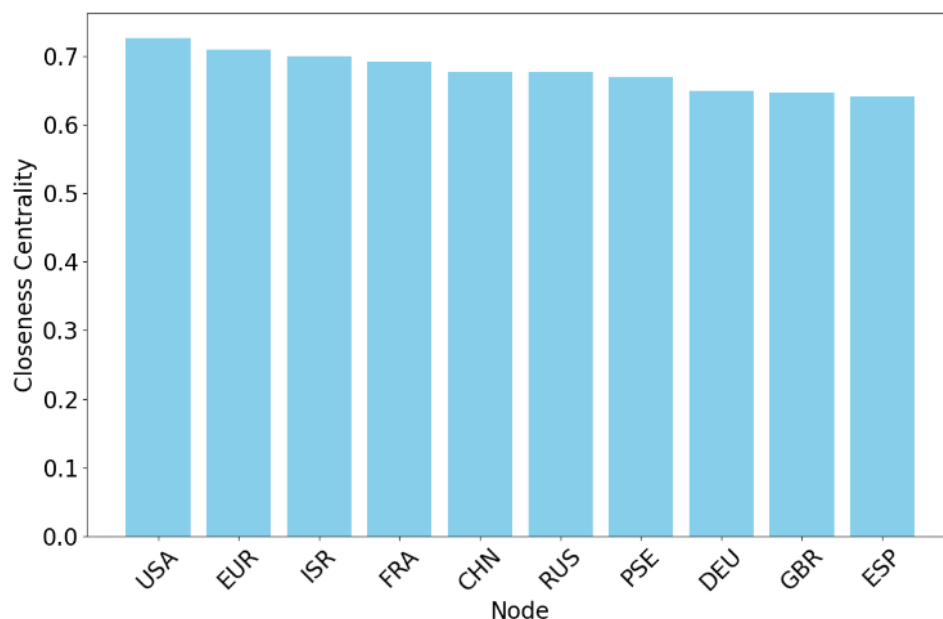


Fig. 3 Top 10 nodes by closeness centrality.

Fig. 3 illustrates the top ten nodes based on their closeness centrality scores. We can find that the differences between each node are relatively small. The top-ranked USA has a value of 0.7262, while the tenth-ranked ESP (Spain) has a value of 0.6409, resulting in a minor difference of 0.0853.

The USA has the highest closeness centrality, indicating it has greater influence and ability to communicate or transmit information in the global network. EUR (Europe) also holds a higher closeness centrality because it represents a collective of multiple countries. Among the top ten countries listed, countries are involved in different events during the selected period, such as wars. Furthermore, the countries shown in the Fig. 3 are not identical to those in Fig. 2. This implies that the importance of nodes is not constant; different methods can produce different important nodes, which also represent various implications.

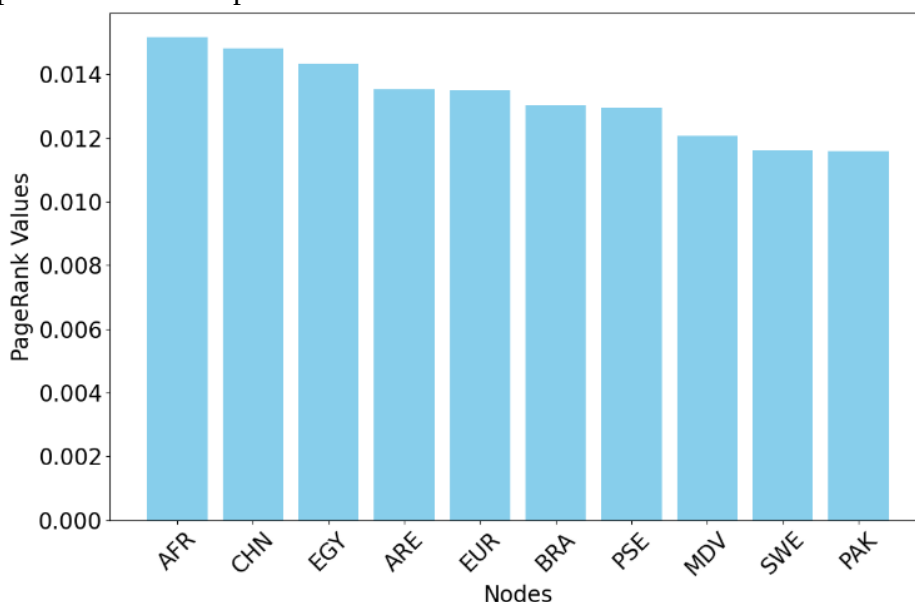


Fig. 4 Top ten nodes by PageRank.

Fig. 4 displays the top ten crucial nodes obtained through PageRank. We can find that the PageRank values of these nodes are relatively close, ranging between 0.011 and 0.014.

AFR (Africa) attains the highest PageRank value, although its degree does not rank within the top ten. This implies that although AFR does not have many connections, those linking to it are more important. Similarly, CHN(China) ranks second, indicating that most countries linked to it hold a

large importance. In addition, ISR was shown in the top ten in degree and closeness centrality but does not appear in the top rank of PageRank, which means the localized conflicts in Israel might not significantly impact the global. The reason could be that Israel is frequently involved in wars, which have become commonplace. As a result, even if global media still pays much attention to the war, it will have less importance.

3.2. State of the country

Table 1. The overall assignment and average assignment for the top ten countries in degree, closeness centrality, and pagerank. the average assignment is the overall assignment divided by its degree.

Node name	Overall assignment result	Average assignment result
	Degree Centrality	
ISR	1865	0.2818
PSE	2122	0.3608
RUS	1286	0.2646
UKR	1448	0.3588
USA	1561	0.4812
EUR	1655	0.5779
DEU	860	0.5212
ESP	783	0.4866
FRA	875	0.5479
EGY	1170	0.7429
	Closeness Centrality	
USA	1561	0.4812
EUR	1655	0.5779
ISR	1865	0.2818
FRA	875	0.5479
CHN	794	0.6208
RUS	1286	0.2646
PSE	2122	0.3608
DEU	860	0.5212
GBR	678	0.4321
ESP	783	0.4866
	PageRank	
AFR	294	0.6100
CHN	794	0.6208
EGY	1170	0.7429
ARE	169	0.7071
EUR	1655	0.5779
BRA	332	0.6348
PSE	2122	0.3608
MDV	73	0.2192
SWE	354	0.4544
PAK	65	0.2802

The above table 1 shows the overall and average assignment results for different countries. We can find that the magnitude of overall assignment results does not correlate with the ordering in degree, closeness, or PageRank. A country with a high degree or closeness centrality score might have a small overall assignment result. Moreover, if we focus on the average assignment result, it does not necessarily correspond proportionally to the overall assignment score. For example, ISR has an

overall assignment of 1865, greater than EUR's 1655. However, ISR's average assignment is 0.2818, less than EUR's 0.5779.

Furthermore, we can observe that all assignment results are positive, indicating that countries are in a state of cooperation that aligns with the global goal of cooperation and mutual benefit in globalization. In addition, countries involved in conflicts or wars, such as ISR and RUS, show a smaller average of assignments. This implies their limited collaboration with other nations in the selected period, likely influenced by the ongoing conflicts.

4. Conclusion

In order to study the importance of a country in the world and each country's corporate tendency, this paper establishes an undirected weighted global network using a part of data selected from The GDELT Project in which each node represents a country.

From the degree distribution, we find that our network is a scale-free network. After that, we explored the core countries from different metrics and analyzed the reasons behind their occurrence. We believe that the occurrence of war is why some countries, such as ISR and UKR, have numerous related events. In addition, by considering the event type, we study the cooperative tendency of a country. Based on our findings, an overall trend of cooperation rather than conflict exists among nations, and countries show different levels of cooperation. This phenomenon could be influenced by globalization.

References

- [1] Fairouz Mustafa Hamdi. The impact of globalization in the developing countries. *Developing Country Studies*, 3(11):142–144, 2013.
- [2] Emilie M Hafner-Burton, Miles Kahler, and Alexander H Montgomery. Network analysis for international relations. *International organization*, 63(3):559–592, 2009.
- [3] G Fagiolo and G Santoni. Country centrality in the international multiplex network. *Netw. Sci.*, 3:377–407, 2015.
- [4] Xin-lei An, Li Zhang, Yin-zhen Li, and Jian-gang Zhang. Synchronization analysis of complex networks with multi-weights and its application in public traffic network. *Physica A: Statistical Mechanics and its Applications*, 412:149–156, 2014.
- [5] Wenwu Yu, Jinde Cao, Guanrong Chen, Jinhu Lu, Jian Han, and Wei Wei. Local synchronization of a complex network model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):230–241, 2008.
- [6] Runjie Xu, Chuanmin Mi, Rafal Mierzwiak, and Runyu Meng. Complex network construction of internet finance risk. *Physica A: Statistical Mechanics and its Applications*, 540:122930, 2020.
- [7] Cristiana J Silva, Guillaume Cantin, Carla Cruz, Rui Fonseca-Pinto, Rui Passadouro, Estevão Soares Dos Santos, and Delfim FM Torres. Complex network model for covid-19: human behavior, pseudo-periodic solutions and multiple epidemic waves. *Journal of mathematical analysis and applications*, 514(2):125171, 2022.
- [8] Michael Hay, Chao Li, Gerome Miklau, and David Jensen. Accurate estimation of the degree distribution of private networks. In *2009 Ninth IEEE International Conference on Data Mining*, pages 169–178. IEEE, 2009.
- [9] Francisco Aparecido Rodrigues. Network centrality: an introduction. *A mathematical modeling approach from nonlinear dynamics to complex systems*, pages 177–196, 2019.
- [10] Junlong Zhang and Yu Luo. Degree centrality, betweenness centrality, and closeness centrality in social network. In *2017 2nd international conference on modelling, simulation and applied mathematics (MSAM2017)*, pages 300–303. Atlantis press, 2017.
- [11] Ying Ding, Erjia Yan, Arthur Frazho, and James Caverlee. Pagerank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, 60(11):2229–2243, 2009.

- [12] Rada Mihalcea. Graph-based ranking algorithms for sentence extraction, applied to text summarization. In Proceedings of the ACL interactive poster and demonstration sessions, pages 170–173, 2004.
- [13] Albert-László Barabási and Eric Bonabeau. Scale-free networks. *Scientific american*, 288(5):60–69, 2003.