Social Network Analysis for Food and Feed security: the European Union case

Alberto Nogales (alberto.nogales@ceiec.es)
Universidad Francisco de Vitoria https://orcid.org/0000-0003-4951-8102
Marçal Mora Cantallops
Rodrigo Díaz Morón
Álvaro García-Tejedor

Research Article

Keywords: Food and Feed Safety, Graph Theory, Social Network Analysis

Posted Date: March 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1457732/v1

Social Network Analysis for Food and Feed security: the European Union case

Graph analysis for European contaminated products

Alberto Nogales[1], Marçal Mora-Cantallops[2], Rodrigo Díaz Morón[1], and Álvaro J. García Tejedor[1]

[1] CEIEC Research Institute, Universidad Francisco de Vitoria, Ctra. M-515 Pozuelo-Majadahonda km. 1,800, 28223 Pozuelo de Alarcón (Madrid), Spain

[2]Computer Science Department, Universidad de Alcalá, 28801 Madrid, Spain

[*]Corresponding Author: Alberto Nogales. Email: alberto.nogales@ceiec.es

Abstract: This paper reports a quantitative and structural analysis of data gathered from the food and feed issues reported by the European Union during the last forty years. This study includes the use of statistical measures and social network analysis techniques. For that purpose, a graph has been constructed considering how the different contaminated products have been distributed alongside countries. This work aims to leverage insights into the structure formed by the involvement of the European countries in the exchange of goods that can cause problems to the population. Results obtained show the roles of the different countries in the detection of sensitive routes. In particular, the analysis shows that there are problematic origin countries, like China or Turkey, whereas European countries, in general, do have good border control policies for the import/export of food and feed.

Keywords: Food and Feed Safety, Graph Theory, Social Network Analysis.

1 Introduction

We live in a globalized world where it is easy to find products from any country in any store. This is an advantage for consumers who, however, are often unaware of the dangers to which they are exposed. Risks arise due to the different policies and legislation that countries must comply with consumer rights and feed and food safety. As a result, Jongwanich (2009) and Timmis (2017) conclude that regulations established by developed countries are holding back food exports from developing countries because of the special mistrust that exists about these products closely related to human health. This situation also affects consumption, as in the case of the 1994 Bovine Spongiform Encephalopathy (BSE) crisis, Verbeke (2001) or cucumbers contaminated with E-coli in Germany during 2011, Bitsch & Rombach (2014).

Since the benefits of food security are directly related to the costs in health systems Pal et al. (2015), there is a growing interest in determining how food security policies in different countries can influence the number and severity of alerts detected in import/export operations. For example, Kleter et al. (2009) reported that European countries had problems with products from China due to the use of banned pesticides on the territory of the European Union (EU).

The Commission of the European Communities considers the highest possible food safety standards to be a key political priority, Mutukumira & Jukes (2003). To this end, the Commission, alongside the EU Member States, European Food Safety authority (EFSA), European Surveillance Authority (ESA), Iceland, Liechtenstein, and Norway, established the Rapid Alert System for Food and Feed[1] (RASFF) in 1979. This tool enables the efficient exchange of information about food and feed issues that can be a risk for human health or can be considered a fraud. The food issues are registered by contact points, which work at a national level, in an online system called RASFF

[1] https://ec.europa.eu/food/safety/rasff/portal_en

Portal[2]. The registration of the notifications is shared between countries and organizations, so they can help to get fast and effective decisions.

The issues recorded in RASFF are made up of a set of categorical characteristics that encode the countries involved and the type of contaminated or hazardous food. They describe the movement of a product between countries and useful insights can be leveraged by taking advantage of the representation of import and export information in the form of a graph or network and its subsequent analysis. The process used to analyze graphs is called Social Network Analysis (SNA). SNA emerged as a branch of application in the field of social sciences to understand how groups interact and behave. It comprises a set of graph theory techniques that can be used to analyze networks formed by various actors. In this case, SNA will be applied to the countries and their trade of products; their behavior can be analyzed as a network. Results will provide food safety patterns that could be used by food policy authorities and official organizations.

This research aims to provide a complete snapshot of the chain trade formed by the involved countries, the issued products, and their hazards, covering the food alerts stored in the RASFF records since its inception. To build the graph, RASFF data representing the countries involved (origin, distribution, and destination), the product issued, and its hazard was used. In this graph, nodes represent countries, while the edges describe the flow followed by food alerts. Edge directions will be set by the role of the country in the commercial chain: origin, distribution, or destination. Edges will also be labelled with the issued product and its hazard. The detailed analysis of the graph allows the extraction of general statistical measures and SNA metrics calculated from the general graph. These metrics can be used to replicate good food policies in countries that have more problems with contaminated products, increase the monitoring on certain products coming from certain countries or finding sensitive routes of trading. This is the first time that the complete RASFF dataset is analyzed as a graph using SNA and a complete set of metrics and statistics are reported and interpreted.

The rest of the paper is structured as follows. Section 2 provides a brief description of the state of the art and background papers. Section 3 describes the materials and methods applied in the study. Section 4 discusses the results obtained and provides a detailed analysis. Finally, conclusions and perspectives for future research are provided in Section 5.

2 Related works

Since RASFF import/export information can be understood as a graph, SNA techniques will be applied to extract and analyze the information contained in this data. SNA is defined in general terms as a way to find patterns of relationships between social entities like people, events, organizations, and other entities Jamali & Abolhassani (2006). Mincer, M & Niewiadomska-Szynkiewicz (2012) gives a more accurate definition, closer to the present research: SNA provides a set of techniques used in graph theory that allow understanding networks formed by several actors. These techniques are often applied, in different fields, to understand the role of different actors in a specific environment and how they interrelate.

In the particular case of food and feed safety, several papers make use of SNA techniques. The model presented in Wu & Guclu (2013) uses SNA in a use case of Maize trade with data from 2000 to 2009. Xu et al. (2014) present a framework for food traceability using data from various resources and applying the Internet of Things, social analytics, and mobile technologies. A network-based methodology for Hungarian cattle holding is presented in Jozwiak et al. (2016). Fair et al. (2017) use information from the global wheat trade between 1986 and 2011 to create a Preferential Attachment network model. Finally, Wang (2018) creates a model for food safety monitoring based on a set of random graphs created artificially. Although these papers use SNA techniques in the field of food and feed safety, they do not take advantage of RASFF data and, thus, are unable to analyze which countries cause more issues or which routes should be monitored more closely due to their risks.

Several papers make use of the RASFF dataset to do statistical analyses, however none of them benefits from the whole historical. In these cases, they are focused on a subset of products, hazards, a particular region or period of time. For example, Dada et al. (2021) studies microbiological hazards in products from the Asia Pacifica region from 2000 to 2020 or Papapanagiotou (2021) analyses Food Contact Materials from 2012 to 2019. Similar works can be found in [Kępińska-Pacelik & Biel (2021),] Alshannaq & Yu (2021), Caldeira et al. (2021), Somorin et al. (2021), De Leo et al. (2021), Djekic et al. (2017), Amico et al. (2018), Kononiuk & Karwowska (2017), Duan et al. (2017), Çınar et

[2] https://webgate.ec.europa.eu/rasff-window/portal/

al. (2017), Czepielewska et al. (2018), Pennone et al. (2018), Papapanagiotou (2017), Van Asselt et al. (2018) and Tudela-Marco et al. (2017)].

There are other papers that apply advanced computer science algorithms to RASFF data. For example, Pigłowski (2021) makes clustering analysis of RASFF in the period of 1999-2018 alongside with data from the Standard International Trade Classification. Pigłowski (2017) also applies cluster analysis to understand the dependencies between products and hazards. In Pigłowski (2017), tree clustering and k-means are applied to RASFF data from 2000 to 2015 to find European dangerous countries. Bouzembrak et al. (2018) applies Bayesian Networks to determine the probability of the contamination in species and herbs.

Only a few papers apply SNA to the particular case of RASFF data. Most of them, however, are very specific. For instance, Petroczi et al. (2011) focuses only on mycotoxin contamination. Bui-Klimke et al. (2014) studies how aflatoxin regulations have affected the worldwide trade of pistachios. Popp et al. (2018) also applies SNA in the field of food safety but it uses a dataset from FAO to study the honey trade network.

Finally, a couple of tools using RASFF dataset have been identified. Naughton et al. (2015) describe a prototype tool to make SNA with RASFF but it lacks metrics or quantitative results, and it is no longer available. Robson et al. (2021) presents a tool to identify threads and risks in the beef supply chain in particular.

Thus, although there are some papers using SNA with RASFF data, none of them works with the whole dataset but Saurkar et al. (2018), which describes the functionalities of a tool without providing SNA metrics. The present work uses all the issues stored in the RASFF portal, instead of a subset, and reports a complete analysis of quantitative results that have not been found in previous literature. Since RASFF information can be modeled as a graph, SNA techniques will be applied to extract and analyze the information in the full RASFF dataset.

3 Materials and methods

This section describes the dataset used to build the graph, and the techniques applied in this work.

3.1 Materials

The information stored in RASFF cannot be downloaded at once as the website only allows users to dump 5,000 records in an XLS file out of a total of 56,351 records that were registered in the RASFF portal from 1979 until the moment of the research. Also, the obtained file does not contain all the registered fields. For instance, it does not include information about distribution and destination countries which are necessary for the described research. In our case we have downloaded all the info necessary to build the graph.

Table 1 Features of a RASFF record used in this research.

Feature	Description
Notification country	Country registering the issue. The cardinality is 32 (EU countries, Norway, Liechtenstein, and Iceland).
Product	The specific name of the product.
Hazard	The hazards or anomalies that have caused the issue.
Origin country	Country of origin of the product. Any country in the world.
Destination country	Country or countries of destination of the product. Any country in the world.
Distribution country	Countries within the transportation chain of the product between origin and destination. Any

country in the world or certain international regulatory organizations.

It must be noted that some issues have more weight in the dataset, thus having an influence in the results. This could be related to stricter food policies in the countries that detect them or because there is a greater number of issues related to a particular product in terms of productivity.

3.1 Methods

For this research, a two-stage analysis has been carried out. First, a statistical analysis on the scrapped RASFF dataset to describe the characteristics of the different elements of the dataset, the countries and the hazards reflected by registered food issues. Results can be used to identify potentially dangerous countries, countries with good food policies or products and hazards that should be closely monitored. Second, an analysis based on social network techniques to understand the graph created from the information of the issues registered in RASFF. This information reflects the number of issues reported among countries: the characteristics of the different actors, their behavior and roles, and the structure of the graph itself.

3.1.1 Statistical analysis

This analysis is based on the frequency analysis of the data contained in the RASFF dataset. The aim is to obtain:

- From the information of countries and routes, the frequency distributions of the most involved countries, considering the role each country plays in the chain: origin, distributor, destination, and notifier. The information will reflect the countries with the most restrictive or effective security policies. Also, sensitive routes will be found. With this information, hypothesis about which countries should increase their food security policies, or which are the countries with good food controls that could be replicated in others could be established.

- Based on the information of the products affected by the reported issues, which products generate the most problems, or which are the most common dangers, as well as the combinations between them. Countries may use this information to pay more attention to specific products or risks or to adjust the toughness of product security policies in certain countries.

3.1.2 Graph SNA

The next step consists of applying SNA techniques to study the graph formed by RASFF issues. A graph G is defined in Schaeffer (2007) as a pair $G=(V, E)$ where V is a set of vertices or nodes, being the number of nodes $n=|V|$ is the order of the graph, and the set E contains the edges of the graph. The edges also have a weight denoted by w that represents the number of occurrences of the relation between the two nodes connected by an edge. We use NetworkX[3], a Python library developed to deal with complex networks/graphs, to create a graph using the data extracted from RASFF.

The first step is to represent the data in the form of a graph $G = (N, E)$ where N is the set of nodes represented by the countries, and E is the set of directed edges that represent the relationship between countries in each RASFF record and the character of the nodes as origin, destination or distribution (this has only been taken into account to stablish the direction of the edge). To include the number of issues between different countries, a weight is added to the edges, reflecting the number of occurrences of an issue between two given countries of a product contaminated by the same hazard. Thus, a weighted Directed Acyclic Graph, $DAG=(N,E,W)$ is obtained, where W is the set of weights of the graph. This is a unique graph that contains all the registered issues in RASFF. In this case, the nodes denote countries, the edge and its direction how a product has traveled through the countries (origin, distribution and destination) and

[3] https://networkx.github.io/

the label of the edge represents the product and the hazard of the issue. Fig. 1 shows how a record is transformed into a part of the graph.

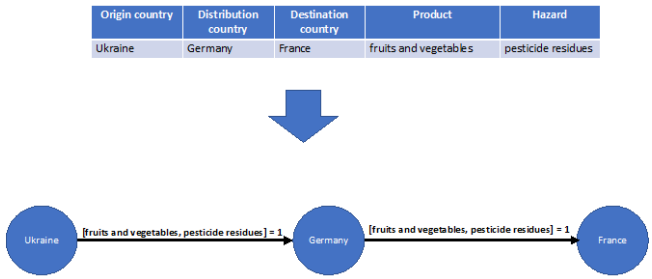

Figure 14: An example of the RASFF record transformed into a part of the graph.

SNA analysis processes RASFF data as a unit, using a DAG formed by all the previously scrapped registered issues, considering the connections created by the transport of a contaminated product through different countries at any date. With such analysis, three types of results are obtained: the general ones, those related to connectivity, and those linked to centrality. General metrics describe some characteristics of the graph itself that will allow to understand the structure formed by the different food issues. Connectivity metrics inform how the different countries are related allows a better understanding of the network performance. Finally, centrality provides information about the relative importance of each country in the network.

3.1.2.1 Definition of general metrics

The density is the number of edges in the network compared to the number of potential edges, Goldberg (1984). This indicates how the nodes are connected between them. It can be used to know if most of the countries can spread a contaminated product to a lot of countries. In equation 1 is defined where m is the number of edges and n is the number of nodes.

$$d = \frac{m}{n*(n-1)} \qquad (1)$$

Determining whether the network is heterogeneous or homogeneous is also relevant. In heterogeneous networks most of the nodes have only a few edges, that is, there are only a few nodes with many edges. This describes how the connectivity of the countries is distributed. In contrast, for homogeneous ones, most of the nodes are not highly connected and have approximately the same number of edges. This kind of network can be modeled by a normal distribution.

When a graph is heterogeneous, its degree distribution follows a power law. Power laws are defined in Clauset et al. (2009), as a model of the relationship between two variables that are inversely related. This is a phenomenon that can be found in the nature where 20% of a population causes 80% of the phenomena. In this particular case, 20% of the nodes should have a high degree. The mathematical definition of a power law is:

$$p(x) = Cx^{-\alpha} \text{ for } x > x_{min} \qquad (2)$$

where x corresponds to the quantity whose distribution is studied, and C and α are constantly related by the expression in Equation 2. Then, x_{min} is a lower bound as the distribution diverges at zero, having a value greater than 0. To normalize the power law, the conditions $\alpha>1$ and $x_{min}>0$ must be fulfilled.

$$C = (\alpha-1)x_{min}^{\alpha-1} \qquad (3)$$

3.1.2.2 Definition of connectivity metrics

Definitions of connectivity metrics can be found at Barnes (1984). A connected component is defined as the biggest subgraph that can be obtained from a graph where all pairs of nodes are connected by a path. By obtaining the connected components, groups of nodes that have a similar behavior between them but differently enough compared to others can be found.

A graph is strongly connected when for every pair of nodes, it is possible two find a path that connects them. If there is a maximal strongly connected subgraph, it can be considered as a Strongly Connected Component (SCC). A graph is weakly connected when turning a directed graph into an undirected, it is strongly connected, being able to find a Weakly Connected Component (WCC). A WCC is a subgraph with some path that can connect all the nodes.

The diameter is related to the type of connectivity of the whole graph. It is defined as the maximum of the shortest paths between all pairs of nodes, West (1996). It is described in the following equation.

$$diam(G) = max_{x,y \in X} d(x,y) \qquad (4)$$

When the diameter is low, it is easier to reach the rest of the network starting from a node. For the RASFF case, it represents the minimum number of connections to reach one country from another in a certain subgraph of the data.

The clustering coefficient of a node is a measure to understand whether a node's neighbors are also linked, Holland & Leinhardt (1971). It is also denoted in Equation 5 where d_n and e_n correspond to the degree of the node and its number of edges.

$$C_n \begin{cases} 0 if d_n = 0 \\ \frac{e_n}{\binom{d_n}{2}} \quad if d_n \geq 2 \end{cases} \qquad (5)$$

When applied to the whole network, it is called the average clustering coefficient. In the present case, it helps to detect whether there are sensitive routes between all the countries, as this measure is useful to understand how connected the nodes are to each other.

3.1.2.3 Definition of centrality metrics

Four different centrality measures will be used: degree, closeness, betweenness, and eigenvector centrality. Degree centrality is related the nodes' degree (that is the number of edges in a node), Freeman (1978). In the case of directed graphs, it is necessary to differentiate between in-degree and out-degree; the first one is the number of edges reaching a node, and the second one is related to the edges leaving the node. This metric can be used to understand which nodes are the most influential and their dependency on the network. In the case of in-degree, it indicates countries that receive more products. Out-degree is used to know the countries that export or distribute the most. Both metrics are defined in Equations 6 and 7.

$$C_j^{OUT} = \sum_{i=1}^{n} a_{ij} \qquad (6)$$

$$C_j^{IN} = \sum_{i=1}^{n} a_{ij} \qquad (7)$$

Closeness centrality is calculated based on the number of edges between a node and each other node, Bavelas (1950). In Equation 8, it is formalized where *n* is the number of nodes and *d(v,u)* is the shortest path between this pair of nodes.

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)} \qquad (8)$$

It gives information about how easy an issued product can be spread through all the countries. When the value is small, spreading information costs less, so the nodes are considered reference points in the network. In this paper, it is useful to know how fast an issued product could reach a lot of countries.

Betweenness centrality measures the importance of a node in connecting other nodes in the network, Bonacich (1972). The following equation describes it. *V* is a set of nodes, $\sigma(s,t)$ is the number of shortest (s,t)-paths and $\sigma(s,t|v)$ is the number of those paths passing through some node *v* over than *s, t*.

$$c_B(v) = \sum_{s,t \in v} \frac{\sigma(s,t \vee v)}{\sigma(s,t)} \qquad (9)$$

It can be used to know the role of a node as an intermediary between nodes that are transmitting the information. Here, it can be measured the importance of the distributor countries.

Eigenvector centrality explains how a node influences the rest of the network. The eigenvector of a

node is high when it is connected to a lot of nodes and they also have a lot of connections, Bihari & Pandia (2015). Equation 10 describes it where $a_{v,t}$ is an element if the adjacency matrix and λ is a constant

$$x_v = \frac{1}{\lambda}\sum_{t\in M(v)} x_t = \frac{1}{\lambda}\sum_{t\in G} a_{v,t}x_t \quad (10)$$

It is useful to rank the nodes in terms of how they influence others in the network. Countries with high eigenvector are very influential to the rest of the network

4 Results and discussion

The results are grouped following the techniques and methodologies described in section 3: statistical analysis of the RASFF dataset and SNA of the graph formed by all the issues.

4.1 Statistical analysis of countries

The general analysis of the countries is conducted based on their roles in the notification system: origin, distribution, destination, and notifier. This information uses the number of times the countries appear in the issues showing which countries perform the most as origin, distribution destination, or notifiers and finding sensitive routes.

Fig. 2 depicts the number of incidents in which the 35 most frequent countries in the issues collected by RASFF have been implicated as origin country. As shown, three countries stand out from the rest: China, Turkey, and India. A few other countries also seem to produce an important number of issues. Nevertheless, most of them are concentrated in the low part of the distribution.

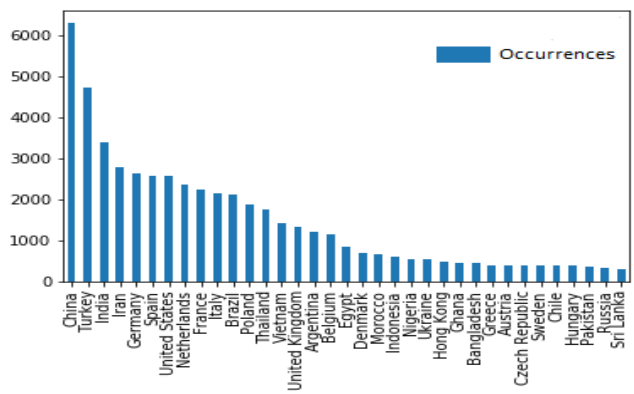

Figure 1: Most frequent origin countries (n=35)

Table 2 shows the top 5 countries that appear as origin country. The table indicates the number of instances per country and its percentage. China is by far the most issued origin country. Also, Turkey and India stand out concerning the others. From the third position onwards, the percentage's differences are less than 1%.

Table 2. Top 5 origin countries.

Country	Instances	Percentage (%)
China	6,280	11.1 %
Turkey	4,707	8.35 %
India	3,375	5.98 %
Iran	2,770	4.91 %

Germany	2,628	4.66 %

Fig. 3 depicts the top 35 distribution countries. The diagram of bars shows how almost all countries belong to the EU. The main difference in Fig. 3 compared to Fig. 2 is the use of the label "No country" which are issues with no distribution country. Then, less countries are found in the medium-range. Most of the countries are distributed in the low part of the graphic.

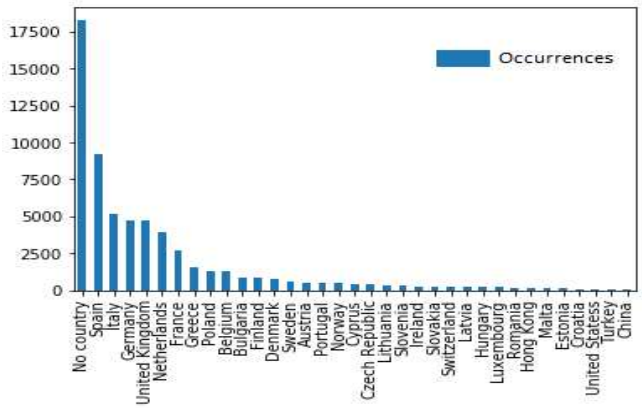

Figure 3: Most frequent distribution countries (n=35)

In Table 3, the top 5 distribution countries with their instances and the percentage they represent are represented. It is worth noting that more than 30% of the reported issues do not have a distribution country. This fact reflects that many of the trades are done directly between two countries (no intermediary country) or that many issues are already reported in the origin country. Another interesting point is that the rest of the countries belong to the EU, which is also plausible, as RASFF is a tool from the European Commission.

Table 3. Top 5 distribution countries.

Country	Instances	Percentage (%)
No country	18,240	32.36 %
Spain	9,202	16.32 %
Italy	5,138	9.11 %
Germany	4,746	8.42 %
United Kingdom	4,716	8.36 %

Fig. 4 shows how the 35 top destination countries are distributed. Again, almost all countries belong to the EU which could be expected due to RASFF's nature. The bar with more occurrences also denotes that a lot of issues have no destination country. No country stands out, even though Germany and Italy appear to be higher. The rest of the distribution is similar to the one in Fig. 3.

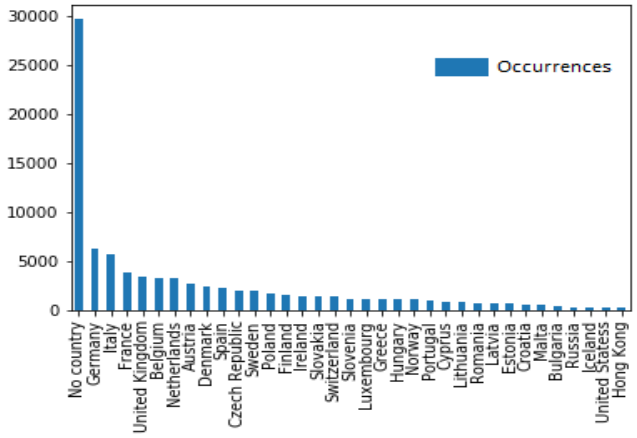

Figure 4: Most frequent destination countries (n=35)

The top 5 destination countries are shown in Table 4, again with instances and percentages. On one hand, it should be noted that there could be more than one destination country in a reported issue. On the other hand, the first position corresponds to issues with no destination country. This occurs when the problem has been found before the product leaves the origin country or at the distribution moment.

Table 4. Top 5 destination countries.

Country	Instances	Percentage (%)
No country	29,612	52.54 %
Germany	6,220	11.03 %
Italy	5,598	9.93 %
France	3,866	6.86 %
United Kingdom	3,372	5.98 %

The information related to the type of countries provided above can be complemented with the one related to the notification countries. These are the ones that have found the hazard in the product. Notification is an action likely to occur in any of the stages during the trading chain. It could be done before leaving the origin country, at the moment of being distributed by another country, or at the destination just before entering a country. For each issue, only one country can be a notifier. This will make possible to find hot spots in the logistic chain.

The distribution of the 35 top notifier countries can be seen in Fig. 5. Again, almost all countries belong to the EU and have a similar behavior to Fig. 2. In this case, three countries can be highlighted: Italy, Germany, and the United Kingdom. Then, a set of a few countries with enough issues to be studied. Lastly, most of the countries only have very few occurrences.

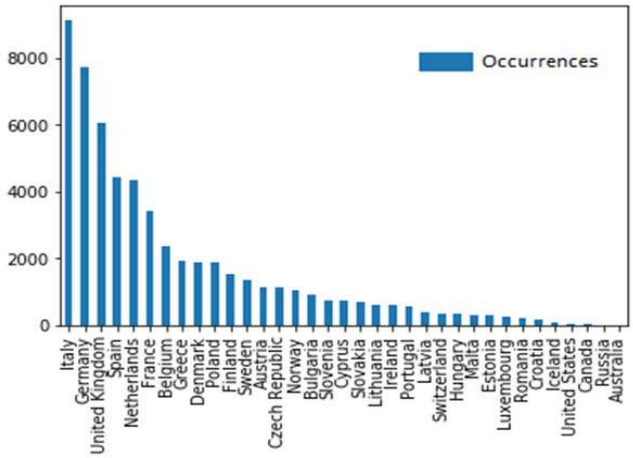

Figure 5: Most frequent notifier countries (n=35)

In Table 5, the top 5 notification countries are shown. As can be seen, except for the Netherlands, the other four countries also appear among the most issued countries by distribution, origin, and destination.

Table 5. Top 5 notification countries.

Country	Instances	Percentage (%)
Italy	9,111	16.16 %
Germany	7,704	13.67 %
United Kingdom	6,061	10.75 %
Spain	4,410	7.82 %
Netherlands	4,357	7.73 %

To find more about the role of notification countries, Fig. 6 relates this role with the others in the trading chain that has been studied above. It shows the percentage of countries that are notifiers and at the same time any of the other roles: origin, distribution, destination, or all of them at the same time (goods that are traded within the country). “Does not apply” are those cases that have an error in the record. The figure shows that most of the time that a country notifies an issue it is at the same time distributor or destination country.

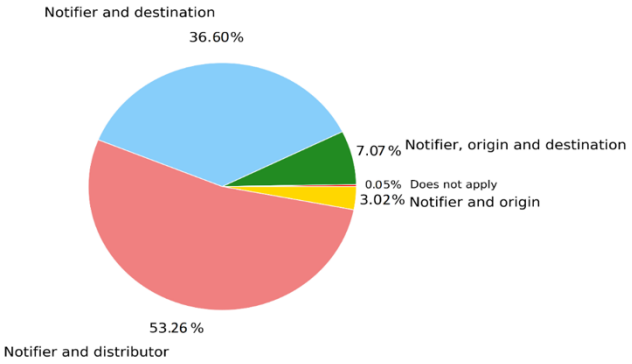

Figure 6: The proportion of countries that are at the same time notifier and any other role at the trading chain.

Table 6 shows the top 5 countries for each case in Fig. 6. The first column represents which are the countries that being countries of origin have notified more issues. The second one provides the same information for countries of distribution. the third column shows the top 5 destination countries that have found contaminated products. The last column contains information on countries that are the origin and destination of a product that has notified the issues. The number between parenthesis is the percentage of issues notified by the country.

Table 6. Top 5 notifier countries vs other roles.

Notifier and origin	Notifier and distribution	Notifier and destination	Notifier, origin and destination
France (256)	Italy (4,781)	Italy (4,154)	Italy (3,553)
Netherlands (247)	United Kingdom (4,127)	Germany (3,619)	Germany (3,006)
Germany (229)	Germany (3,853)	United Kingdom (1,773)	United Kingdom (1,414)
Belgium (215)	Spain (3,695)	France (1,626)	Denmark (1,264)
Italy (171)	Netherlands (2,818)	Denmark (1,394)	France (1,151)

It is also interesting to know from which countries come the contaminated products that the notifiers have found. Two chord diagrams are used to represent the routes followed by the reported products, shown in Figure 7 and Figure 8. The first accounts for countries that are notifiers and distributors, while the second refers to countries that are notifiers and destination at the same time. In both diagrams, countries from Table 6 are in the upper half of the circle. Then, they are connected to the origin countries of a registered issue by a specific colored chord. The width of the chord depends on the number of issues found having one country as the origin/destination and the other as a notifier.

In Figure 7, we can see how there are three country pairs with a wider connection in comparison with the rest of the considered countries. In particular China-Italy, Iran-Germany, and India-United Kingdom.

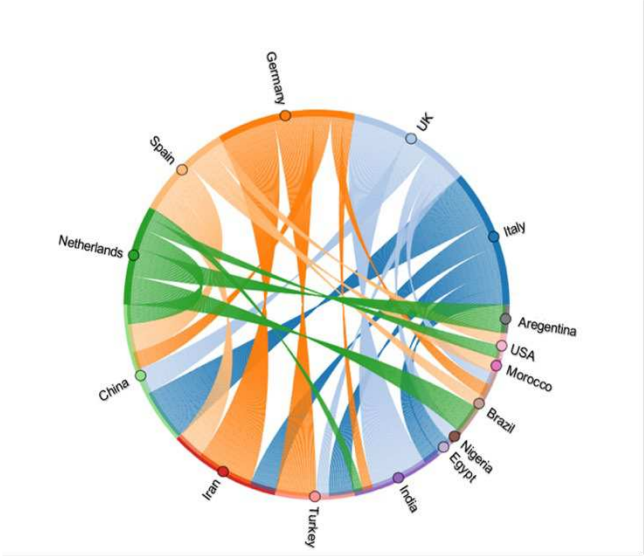

Figure 7: Chord diagram. Top 5 notifiers and distributor countries with their origin countries.

Fig. 8 shows relations between destination and notifier countries, some cases can be highlighted. The biggest connection is from Spain to Italy. Then, there is a group that almost has the same size: Germany-Netherlands, Germany-Turkey, or Germany-China. As can be appreciated in this graphic all the destination countries are from Europe.

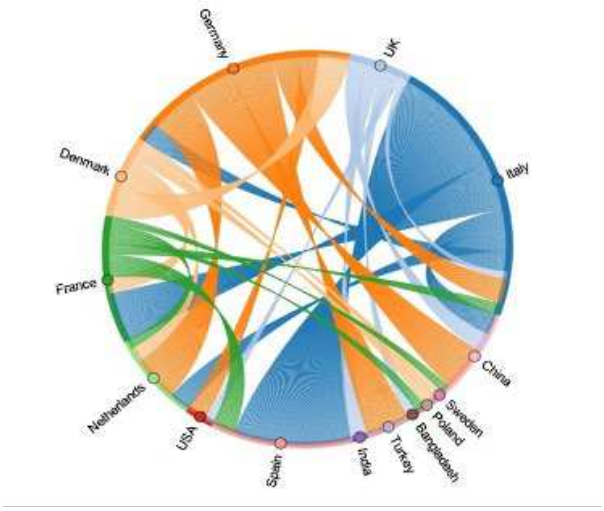

Figure 8: Chord diagram. Top 5 notifiers and destination countries with its origin countries.

4.2 Statistical analysis of issued products

The following analysis focuses on which products are most frequently contaminated and which contaminants are most common. This information is useful, so organisms could pay more attention to these products or increase the analysis to detect these hazards.

Table 7 shows the most issued products where the columns correspond to the number of instances and their percentage concerning the total. Nuts are by far the products that generate more problems. Problems in fruits and vegetables, and fish and fish products are also common. The rest of the products don't seem

to give as many problems.

Table 7. Top 5 issued products.

Product	Instances	Percentage (%)
Nuts, nut products, and seeds	10,454	18.55
Fruits and vegetables	8,321	14.76
Fish and fish products	5,956	10.56
Meat and meat products (other than poultry)	3,067	5.44
Food contact materials	2,910	5.16

Table 8 shows the 5 most frequent hazard that caused a problem in the product. Mycotoxins (toxic metabolites produced by the fungus kingdom) being the most reported hazard is expected since we have already seen in Table 7 that the most reported products are nuts and fruits, Smith et al. (2016). Pathogenic micro-organisms also appear as the cause of many registered issues as these bacteria can be found in every kind of food, Pigłowski (2019). The problem with pesticides is that depending on the countries they are banned or not, Carvalho (2006), so they can be used with products in one country that exports them later to another country where they are forbidden. Considering Pigłowski (2018), metals are related to fish and food contact materials which are also reflected in Table 7 where fish appears among the most issued products. Finally, microbial contaminants have been reported in different products as processed foods or raw agricultural products, Jha (2015).

Table 8. Top 5 found hazards.

Hazard	Instances	Percentage (%)
Mycotoxins	12,068	21.41
Pathogenic micro-organisms	9,629	17.08
Pesticide residues	6,459	11.46
Metals	4,666	8.28
Microbial contaminants (other)	4,490	7.96

Finally, to have a better understanding of the relation between products and their hazards, Fig. 9 is provided. This graph depicts a bubble diagram of the top products and hazards more registered in RASFF. To obtain this information, all the combinations and their frequency have been obtained. Then, the 10 products with more instances have been drawn in the X-axis while the Y-axis depicts all the hazards related to these products. The size of the bubbles corresponds to the frequency of the combinations and the color is univocal. As can be seen, two issues stand out from the rest. Nut products and seeds are contaminated by mycotoxins which fit with Tables 7 and 8, as they are in the first positions. Also, fruits and vegetables where pesticides have been found and meat products having pathogenic micro-organisms which could also be expected as they are in the high positions of these Tables.

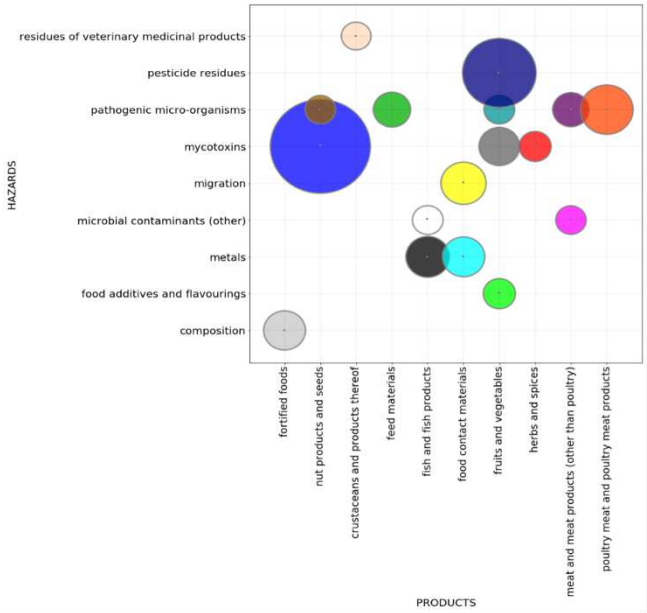

Figure 9: Diagram bubble that relates issued products with its hazard.

4.3 RASFF graph SNA

The dataset has been used to build a graph taking into account the different roles of the countries in the trading chain. A representation of the graph with RASFF information has been obtained using Basemap Matplotlib Toolkit[4] which allows plotting 2D data in maps (Fig. 10). In this Figure, the size of a node represents the total number of issues that concern a country and the thickness of an edge, the total number of issues involving two countries. The color of a node represents the country's role as a mixture of importing (blue) and exporting country (red). The intermediate colors (green) represent countries with a mixed role. The figure shows China as the biggest exporter of issued products, Spain and Germany as importers finding hazards, and France or Poland as balanced countries in terms of finding contaminated imports and distributing or exporting hazardous food.

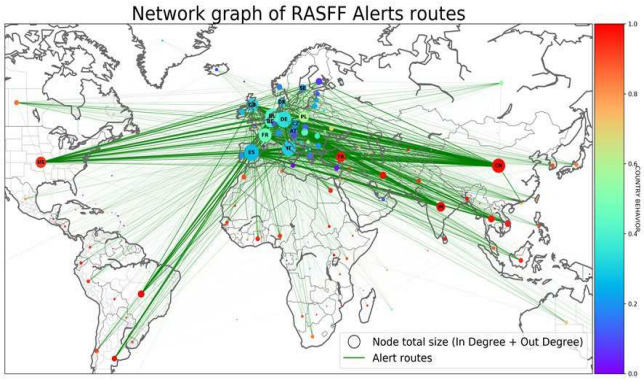

Figure 10: Graph with all RASFF issues represented on a worldwide map.

Figure 11 shows the subgraph of Figure 10 for Europe since the RASFF data comes from the EU countries. As can be seen in Fig. 11, Spain, the United Kingdom, Germany, and Italy are the countries receiving more issued products. Also, Turkey is a sensitive country for exporting products. These trends are consistent with the results obtained in the statistical study previously described in subsection 4.1.

[4] https://matplotlib.org/basemap/

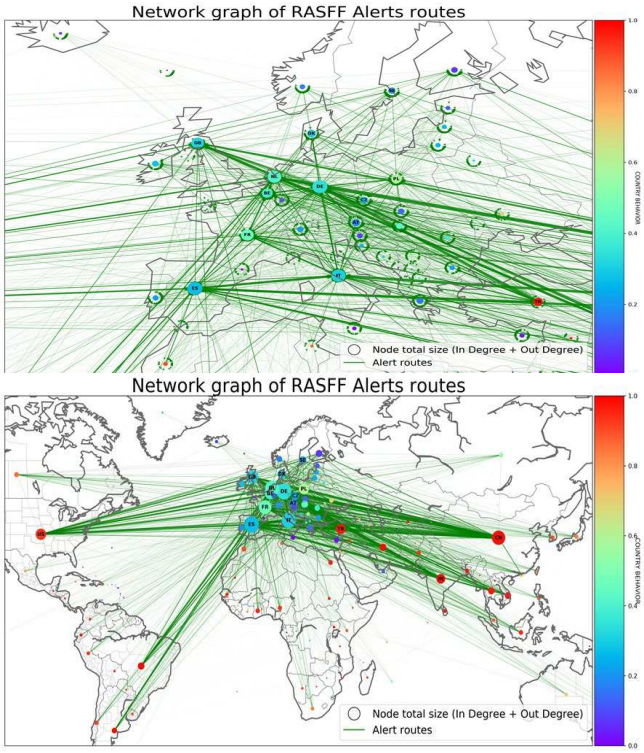

Figure 11: A subgraph of Figure 10 for EU countries.

4.3.1 General SNA metrics

Table 9 shows the metrics for the RASFF graph depicted in Fig. 10. For all metrics, not only the global values are computed (considering the full RASFF graph) but also the metrics for each type of issue sub-category (alert, border detention and information). The number of nodes that correspond to countries in the global network is 222. However, the number of countries in the World are fewer. This mismatch is explained by two factors: territories such as Monaco and extinct countries as the former Yugoslavia can also be found in the historical graph.

Table 9. General metrics.

Metrics	Global	Alerts	Border rejection	Info.
Number of nodes	222	215	164	201
Number of edges	7,092	5,932	1,260	4,010
Density	0.14	0.13	0.05	0.10
Type of network	Heterogeneous	Heterogeneous	Heterogeneous	Heterogeneous

The number of edges that correspond to the total number of products' movements is 7,092 and the RASFF graph density is 0.14.

To study the heterogeneity of the RASFF network, the distribution of the nodes' degree versus the number of nodes having this degree has been analyzed. It can be seen in Fig. 12 that the graph seems to follow a power law. In this plot, a small number of instances are clustered in the left part which is a signal of a heterogeneous graph.

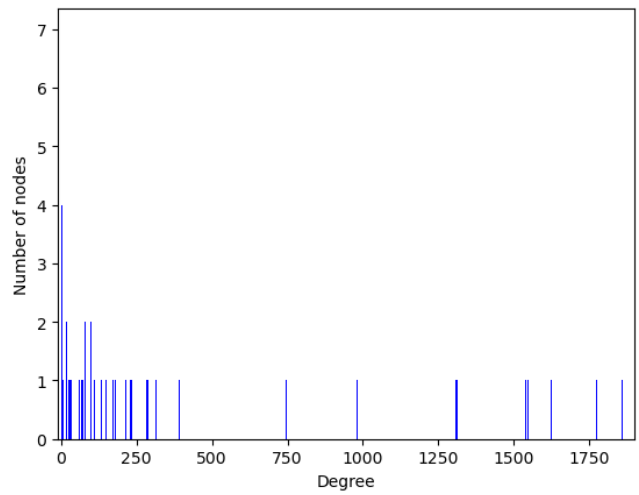

Figure 12: Distribution of nodes' degree and their frequency.

The confirmation has been obtained using a Python library called power-law implemented in Alstott & Bullmore (2014). By using this library, it has been calculated the values for x_{min} and α. The obtained values are 79 and 1.45 respectively. Since α is greater than 1, it can be concluded that the distribution follows a power law. Fig. 13 shows how the distribution fits with the power-law, concluding that the RASFF graph is a heterogeneous network.

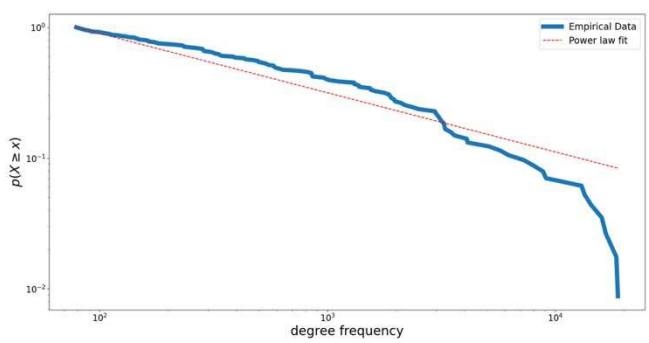

Figure 13: Degree distribution fitting the power law.

4.3.2 Connectivity

The connectivity has also been studied. Table 10 compiles these metrics that can be used to understand how the different countries are connected due to the registered issues. The RASFF graph is a Weakly connected graph with one connected component that contains its 222 nodes. As the graph is not strongly connected, the diameter does not apply. The same results are found in each sub-category.

Table 10. Connectivity metrics.

Metrics	Global	Alerts	Border Rejection	Info.
Strongly	False	False	False	False

connected				
Weakly connected	True	True	True	True
Number of WCC	1	1	1	1
Diameter	Does not apply	Does not apply	Does not apply	Does not apply
Average clustering coefficient	0.78	0.79	0.24	0.77

However, the clustering coefficient should be analyzed in-depth to understand the role of a country as a possible spread vector for possible issues. The rationale behind these results is that there are countries having clustering coefficient 0 and others having 1. We will study in detail the 5 countries with the highest and lowest clustering degree excluding 0 and 1, Tables 11 and 12 respectively.

Table 11 shows countries having a low clustering small non-zero clustering coefficient.

Table 11. Top 5 countries with the lowest clustering coefficient.

Global	Alerts	Border rejection	Info.
France (0.20)	France (0.21)	Italy (0.05)	Netherlands (0.21)
United Kingdom (0.20)	Netherlands (0.23)	United Kingdom (0.05)	Spain (0.22)
Netherlands (0.21)	Belgium (0.23)	Spain (0.06)	Italy (0.23)
Spain (0.22)	United Kingdom (0.28)	France (0.07)	France (0.24)
Italy (0.23)	Italy (0.28)	Malta (0.07)	Germany (0.24)

In Table 12, countries have a high clustering but that does not reach 1.

Table 12. Top 5 countries with the highest clustering coefficient

Global	Alerts	Border rejection	Info.
French Polynesia	French	Suriname	Syria

(0.99)	Polynesia (0.99)	(0.67)	(0.99)
Yemen (0.99)	Burundi (0.99)	Paraguay (0.67)	Madagascar (0.99)
West Bank and Gaza Strip (0.98)	Togo (0.99)	Togo (0.53)	Paraguay (0.99)
Honduras (0.98)	Belize (0.99)	Yemen (0.5)	Ethiopia (0.98)
Burundi (0.98)	Gabon (0.99)	Honduras (0.5)	Sudan (0.98)

4.3.3 Centrality

The last set of metrics is related to centrality. In-degree and out-degree centrality shows how countries behave. In Table 13a and 13b, the values of the top 5 countries of both metrics have been obtained. By studying the in/out-centrality of the nodes, how each country within the network works as a distribution gateway for third party products or whether they represent a focus of problems as countries of origin can be assessed.

Table 13a. Top 5 countries by in-degree centrality

In-degree centrality	Alerts	Border Rejection	Info.
Spain (0.71)	Netherlands (0.68)	Spain (0.67)	Spain (0.71)
France (0.65)	France (0.66)	United Kingdom (0.57)	Italy (0.675)
Italy (0.65)	Belgium (0.64)	Italy (0.53)	Netherlands (0.66)
United Kingdom (0.64)	Spain (0.63)	France (0.42)	United Kingdom (0.65)
Germany (0.60)	Germany (0.62)	Netherlands (0.40)	Germany (0.64)

Table 13b. Top 5 countries by out-degree centrality

Global	Alerts	Border Rejection	Info.
France (0.72)	France (0.76)	China (0.31)	Netherlands (0.59)

Netherlands (0.72)	Belgium (0.72)	India (0.26)	China (0.52)
Belgium (0.68)	Netherlands (0.69)	Turkey (0.21)	France (0.52)
United Kingdom (0.65)	United Kingdom (0.66)	United States (0.18)	Germany (0.51)
China (0.62)	Italy (0.64)	Egypt (0.18)	Spain (0.48)

In Table 14, the top 5 countries by closeness centrality are shown. It measures how close a node is to the rest of the nodes in the graph.

Table 14. Top 5 countries by closeness centrality

Global	Alerts	Border Rejection	Info.
Spain (0.73)	Netherlands (0.74)	Spain (0.69)	Spain (0.74)
France (0.68)	France (0.73)	Netherlands (0.57)	Italy (0.72)
Italy (0.68)	Belgium (0.71)	Italy (0.52)	Netherlands (0.71)
United Kingdom (0.68)	Spain (0.71)	United Kingdom (0.52)	United Kingdom (0.70)
Germany (0.65)	Germany (0.70)	Germany (0.52)	Germany (0.70)

The values of betweenness centrality are gathered in Table 15 and are all very low (close to zero). This metric shows which are the nodes that behave as connectors in a network; apparently, there are no critical brokers.

Table 15. Top 5 countries by betweenness centrality

Global	Alerts	Border Rejection	Info.
United Kingdom (0.07)	France (0.13)	Turkey (0.15)	Netherlands (0.12)
Belgium (0.06)	Netherlands (0.10)	Portugal (0.12)	Spain (0.10)

France (0.05)	Belgium (0.09)	Morocco (0.12)	United Kingdom (0.09)
Italy (0.05)	United Kingdom (0.07)	China (0.12)	France (0.08)
Netherlands (0.05)	Italy (0.06)	Spain (0.10)	Italy (0.08)

The information of the eigenvector centrality for the top 5 countries has been obtained and is shown in Table 16. The measures obtained are near the average value or even lower (eigenvector centrality goes from 0 to 1). This metric gives information about how a node influences the rest of the network.

Table 16. Top 5 countries by eigenvector centrality

Global	Alerts	Border Rejection	Info.
Germany (0.41)	Netherlands (0.17)	Spain (0.48)	Italy (0.19)
Italy (0.41)	Spain (0.17)	Bulgaria (0.39)	Spain (0.19)
France (0.28)	France (0.16)	Netherlands (0.36)	Netherlands (0.19)
Netherlands (0.26)	Belgium (0.16)	Poland (0.32)	Germany (0.19)
Spain (0.26)	Germany (0.16)	Germany (0.29)	United Kingdom (0.18)

5 Discussion

From the results compiled in the previous section, we can conclude several things.

Data related to the origin countries shows the countries recurrent problems with their products or a more intense export activity comparing with others. This indicates that countries should pay attention to the products imported from these places. It should be stated that most of the hazardous products come from Asia (China, Turkey, India and Iran) which is confirmed by Pigłowski (2020). However, as most of the issues are categorized as border rejections, this is problem seems to be under control because these products have not entered the country. Germany is the only EU member, this should be study in-depth as it could be a consequence of detecting issues in products that come from the own country, so their food policies could be taken as an example. Regarding Pigłowski (2017), Germany has a lot of notifications in origin, maybe, due to some problems with meat products. As this type of issues could be a problem to the free market of the European territory without possibilities of border rejections, controls in market should be reinforced for this product category.

From the data of distribution countries, we can see that most of the issues do not have a distributor. Also, we found a set of European countries that have good control policies when distributing products. It

seems that food quality policies with products entering Europe are high. Spain appears above the rest of the countries, probably due to a large number of ports and its central role as a hub of the distribution chain. This is confirmed by the centrality studies and by Caldeira (2021) that describes Spain as the main gateway for fishery products in the EU.

The interesting point in the destination countries table is that three of the countries are the same as in Table 3 (distribution countries). The reason could be that these countries are the ones that import more products, or that they are very concerned about foreign products entering their borders.

Summarizing the study regarding the countries' role in the trading chain two main conclusions are obtained. First, most of the origin countries are not from the EU, maybe due to the different food policies between countries. Second, Germany, Italy, and the United Kingdom are on the top of distributors and destinations, what makes them more active in the trading chain than the rest.

Considering the information of notifier countries, it can be concluded that these countries notify more than others because they their food security policies are stronger. This makes sense by comparing results from Table 5 with Tables 3 and 4. In this case Spain, Italy, Germany, and UK are in the top if distribution and destination demonstrating their high activity, which together with the information from notifier countries make them good in finding issued products.

If we looked at the pie chart of notifier countries, there are two possible interpretations. First, countries receive more products that they export so there are more possibilities in finding these issues or food policies in countries from the UE (most of the countries are member states as seen in Tables 3 and 4) are more restrictive. Second, countries are more concerned about the products they import as food policies in other countries could be different.

An in-depth analysis of the notifier countries and the other roles is in Table 6. First, Germany and Italy are in the four tops. This can be interpreted as these countries have good food safety controls. Analyzing each column separately and comparing the information with Tables 2, 3, and 4, the following points can be concluded. Germany can be considered to be aware of not exporting issued products as it is also at the top of Table 2 (origin countries). By comparing with Table 3 (distribution countries), it is found that countries in the second column are concerned about not distributing contaminated products to others. The same conclusions, related to imported products, can be obtained after comparing the third column with Table 4 (destination countries), except Denmark. The final column depicts countries that are finding food problems internally, so their food policies should be replicated.

Regarding the diagram chords relating countries (Fig. 8 and Fig. 9). In Fig. 8, countries maintain a trade relationship but have different food policies, so these are sensitive routes. It should be remarked that India is a former colony of the United Kingdom, being normal a regular trading between them. Mukherjee et al. (2019) points out that even though India is an interesting market for United Kingdom, the low usage of technology in the food safety field entails problems. So, we proposed that countries with possibilities of stablishing profitable food trading chains should create funding programs to introduce new technologies that help in food safety. In Fig. 9. it can be remarked that the route between Spain and Italy should be revised, and that Germany has a very restrictive food safety policy or is importing a lot of products.

In relation with the results obtained with products and hazards, we can make the following proposals. The appearance of mycotoxins in nuts is difficult to prevent although there are recommendations during all the harvesting cycle. However, as they are directly related to the weather, predictive models in regions with similar weather can be implemented and tested. The decrease of mycotoxins' issues is something that should be addresses as it is directly related with the number of trades Nes & Schaefer (2018). The appearance of pathogenic microorganisms in meat products is reported by Piglowski (2021) which shows a high correlation between the number of kilograms traded among EU countries and the alerts reported by them. As, this happens in a free distribution the market, the best way to reduce these issues is to improve the traceability policies for these products. Finally, there are a lot of issues of fruits and vegetables containing pesticides. In this case, we consider that new and effective Integrated Pest Management (IPM) processes should implemented.

Table 9 shows an important figure that has consequences in the rest of the metrics. Number of border rejections are significantly lower than other type of alerts. With this information, we can conclude that the products with high risk for the population are not so many.

Having a density close to zero means that in the case of a food crisis, it is difficult to distribute that product through a lot of countries. In other words, the trading between all the countries' network is low.

The heterogeneity of the network implies that there are a few countries that dominate the network as they have a lot of connections with the rest. These countries should be studied in more detail as they are the ones generating more problems or the ones that have better food policies that should be replicated.

The average clustering coefficient denotes that when a product leaves its origin how the distribution or destination countries are connected between them. This could lead to a fast expansion of the product in particular parts of the graph.

In this research, we find two groups of countries with clustering coefficient 0 and 1. The first group are non-EU countries that do not make a lot of exports and their destination countries do not trade between them. In the second group, we find countries that seem to not make a lot of exports but in this case, their distribution or destination countries are highly connected between them. Apart from these groups, countries with low clustering coefficient trade with a lot of countries, the possibilities of them being connected are lower. Finally, there are countries with high clustering coefficient. As the network is very large, countries that do not export a lot (no matter the reason) have fewer friends. Then, due to the small number of friends, there are high possibilities that at a particular moment all the trading links between them could be established.

Summarizing the analysis of the clustering coefficient. Countries can be classified into three groups. The first group whose clustering coefficient is zero. They seem to be countries with low exports whose distribution or destination countries are not connected between them. In this case, a food issue will not be spreading through other countries. The second group comprises countries that trade a lot but whose neighbors are not highly connected. In this case, the problems are the same as in the previous use case. Finally, countries with high coefficient clustering seem to be small ones or those whose exports are no high. In these cases, having a food issue is not very common but could be spread to a few countries.

Looking at the in-degree centrality, we see how EU countries, like Spain, appear as difficult-to-save ports in terms of food policy, reflecting the harshness or strict character of their safety standards. In terms of out-degree centrality, the EU countries that appear are those with more relaxed regulations for exporting products. It should be noted that the UK and France, in turn, also appear as countries that impose tough standards on the products they receive, possibly reflecting an imbalance in their food safety standards. In the case of out-degree, China is the non-European country that has more problems with its exports, so Europe must have more control over its products. It also should be noted that all the countries in the border rejection column are non-European countries.

Closeness centrality information can be interpreted as the countries that should be considered as major sources of easy spread of contaminated products. This means that these countries should implement good food policies as they are sensitive points in the network.

Regarding betweenness centrality, it can be concluded that none of the countries act like that in the graph. It means that there are no countries whose removal could cause a disconnection of the network. In terms of issued products, there are no main distributors in the chain.

If we looked at the eigenvector centrality, this can be considered the most important countries having this role. However, as the values are under 0.5, it can be considered that countries are important but not influencing a lot the rest of the graph. In the RASFF network, it supposes that the sensitive routes are not critical.

Two other separate analyses were also conducted:

- Distinguishing between the three main types of issues: notification, alert and border detention.

- Analyzing each decade separately.

In the first case, the results for each type of issue are consistent with the global ones; in some metrics relative differences have been found (e.g., some countries seldom appear among the top countries or change in relative position) but for the most part the same set of UE countries plus has been found to monopolize the network measures (the table considering the clustering coefficient is the only exception). Other isolated cases could be highlighted: Belgium always appears on the top of alerts, and there are some exceptions like Bulgaria, Malta or Morocco. These cases can be studied in-depth as future works. On the second case, the only relevant observation (though arguably obvious) is that the number of alerts grows decade by decade, which only shows the increase in the adoption of the RASFF alert system by trade countries.

6 Discussion

This paper presents a statistical and structural analysis of historical information of food issues previously obtained by scraping the online RASFF portal. First, a statistical analysis of the raw data has been carried out to obtain some significant metrics that are shown in tables and graphics. We then constructed a graph from the most relevant characteristics of each record: origin country, distribution country, destination country, product, and hazard. The result is a weighted DAG where nodes represent the countries and edges describe how the trades flow and the information of the issues (product plus hazard). The direction of each edge is given by the roles of the countries which describe how the products are transported (origin, distribution, and destination country). Weights are computed by summing up the total number of times that an issue occurs involving the same pair of countries. Using this graph, we can characterize the chain trade formed by the countries involved, the issued products, and their hazards by applying SNA techniques. The result is several descriptive metrics of the graph, such as general, connectivity, and centrality metrics.

General metrics (number of nodes, number of edges, diameter, density, and type of network) describe the graph itself. From the study, we may conclude that China and Turkey are sensitive countries as the products imported from those countries produce a lot of issues. European countries' policies seem to be good and consistent as they report a lot of issues when they act as distributors. Most of these issues have no destination registered as they have been found before reaching their expected destination. In this sense, Italy, Germany, the United Kingdom, or Spain notify more cases and could be considered safe countries. From the analysis of the role played by the different countries (origin, distribution, or destination) when reporting a problem (notifier), it can be concluded that European countries are more stringent with food controls than non-European countries. In particular, Germany and Italy can be highlighted as they are always on the top. Regarding hazards and products, nuts, nut products, and seeds contaminated with mycotoxins and fruits and vegetables with pesticides are the ones most issued.

Connectivity metrics are used to check if the network is strongly or weekly connected, its number of connected components, and the average clustering coefficient. It has been checked that RASFF graph has a low density and is a heterogeneous network. The former means that when a food crisis happens, it is difficult that it spreads through the rest of the countries. The latter allows concluding that only a few countries are really important in the connectivity of the network. Related to connectivity, it has been obtained that the graph is weakly connected, having only one component. This means that all the countries can be reached from the rest of the graph. The clustering coefficient has also been analyzed obtaining an average clustering of 0.78 and grouping the countries in four different groups depending on the clustering coefficient value. The average clustering coefficient indicates that an issued product could spread quickly through several countries that are near between them. The different clustering coefficients show how the countries behave in the network. For example, countries with the highest clustering coefficients have the most restrictive border policies.

In-degree and out-degree have been obtained, standing out Spain and France. In the case of Spain, that means that Spain receives a lot of goods (maybe due to the big number of ports) or have good food policies. In the case of France, a high out-degree means that they have some problems with product leaving the country. Regarding centralities, different metrics have been calculated. The closeness value for the graph

is high, indicating that countries at the top should have good protocols as issues produced in their territories can be spread easily. Betweenness has low values, which implies that there are no countries that have greater control over the rest. Eigenvector shows no extreme values, so none of the countries has more contribution to the rest of the network.

As future work, an in-depth analysis of the countries that have been highlighted during the study seems to be necessary. Future work should look at the specific reasons why a country stands: it might be because they import more products or because they have safe food policies. The behaviour of trading products should also be studied as time series. An in-depth study of seasonality could then be carried out with this time series analysis to confirm those seasonal products that are related to the month of the year and the hemisphere where they are cropped. This would provide information on how the countries' behaviour change depending on the season and the products that are trading at this time. For instance, fig issues are mostly from Turkey during September. The isolated study of networks formed by the countries and their trade will be useful when obtaining metrics like clustering coefficient to measure the risk of propagating contaminated products. Another interesting study would consist in comparing the results with the world trade network. This will allow to have a better understood of the conclusions obtained in this paper as we could compare the number of issues with the total amount of exports and imports. Also, making a socioeconomic study of the relations obtained in this paper could potentially be used to propose solutions to the origin of some problems that arose here. Finally, we can benefit from using Graph Convolutional Neural Networks, an artificial intelligence technique in the field of Deep Learning which is currently one the most interesting techniques to manage big amount of data.

Conflict of interest

The authors declared that they have no conflict of interest.

References

[1] Alshannaq, A., & Yu, J. H. (2021). Analysis of EU Rapid Alert System (RASFF) Notifications for Aflatoxins in Exported US Food and Feed Products for 2010–2019. *Toxins*, *13*(2), 90.

[2] Alstott, J., & Bullmore, D. P. (2014). power-law: a Python package for analysis of heavy-tailed distributions. PloS one, 9(1).

[3] Amico, P. D., Nucera, D., Guardone, L., Mariotti, M., Nuvoloni, R., & Armani, A. (2018). Seafood products notifications in the EU Rapid Alert System for Food and Feed (RASFF) database: Data analysis during the period 2011–2015. *Food Control*, *93*, 241-250.

[4] Barnes, J. A., & Harary, F. (1983). Graph theory in network analysis. *Social networks*, *5*(2), 235-244.

[5] Bavelas, A. (1950). Communication patterns in task-oriented groups. *The journal of the acoustical society of America*, *22*(6), 725-730.

[6] Bihari, A., & Pandia, M. K. (2015, February). Eigenvector centrality and its application in research professionals' relationship network. In *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)* (pp. 510-514). IEEE

[7] Bitsch, V., & Rombach, M. (2014). E. coli and fresh produce: comparing media coverage and market effects in the US and Germany.

[8] Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of mathematical sociology*, *2*(1), 113-120.

[9] Bouzembrak, Y., Camenzuli, L., Janssen, E., & Van der Fels-Klerx, H. J. (2018). Application of Bayesian Networks in the development of herbs and spices sampling monitoring system. *Food Control*, *83*, 38-44.

[10] Bui-Klimke, T. R., Guclu, H., Kensler, T. W., Yuan, J. M., & Wu, F. (2014). Aflatoxin regulations and global pistachio trade: insights from social network analysis. *PloS one*, *9*(3), e92149.

[11] Caldeira, A. J. R., Alves, C. P. P., & Santos, M. J. (2021). Anisakis notification in fish: an assessment of the cases reported in the European Union Rapid Alert System for Food and Feed (RASFF) database. *Food Control*, *124*, 107913.

[12] Carvalho, F. P. (2006). Agriculture, pesticides, food security and food safety. *Environmental science & policy*, *9*(7-8), 685-692.

[13] Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. SIAM review, 51(4), 661-703.

[14] Czepielewska, E., Makarewicz-Wujec, M., Różewski, F., Wojtasik, E., & Kozłowska-Wojciechowska, M. (2018). Drug adulteration of food supplements: A threat to public health in the European Union?. *Regulatory Toxicology and Pharmacology*, *97*, 98-102.
[15] Çınar, S., Yılmaz, S. N., Aydın, E., & Yorulmaz, A. (2017). Rapid Alert System for Food and Feed (RASFF) 2009-2016 Turkey Report. *Turkish Journal of Agriculture-Food Science and Technology*, *5*(8), 873-882.
[16] Dada, A. C., Somorin, Y. M., Ateba, C. N., Onyeaka, H., Anyogu, A., Kasan, N. A., & Odeyemi, O. A. (2021). Microbiological hazards associated with food products imported from the Asia-Pacific region based on analysis of the rapid alert system for food and feed (RASFF) notifications. *Food Control*, *129*, 108243.
[17] De Leo, F., Coluccia, B., Miglietta, P. P., & Serio, F. (2021). Food contact materials recalls and international trade relations: An analysis of the nexus between RASFF notifications and product origin. *Food Control*, *120*, 107518.
[18] Djekic, I., Jankovic, D., & Rajkovic, A. (2017). Analysis of foreign bodies present in European food using data from Rapid Alert System for Food and Feed (RASFF). *Food Control*, *79*, 143-149.
[19] Duan, M., Li, Q., Liu, W., Hao, J., Liu, P., Han, B., & Wang, Y. (2017). Analysis on the safety of metallic food contact materials and products in China based on RASFF. *Journal of Food Safety and Quality*, *8*(12), 4865-4869.
[20] Fair, K. R., Bauch, C. T., & Anand, M. (2017). Dynamics of the global wheat trade network and resilience to shocks. *Scientific reports*, *7*(1), 7177.
[21] Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, *1*(3), 215-239.
[22] Goldberg, A. V. (1984). *Finding a maximum density subgraph*. Berkeley: University of California.
[23] Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative group studies*, *2*(2), 107-124.
[24] Jamali, M., & Abolhassani, H. (2006, December). Different aspects of social network analysis. In 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings) (WI'06) (pp. 66-72). IEEE.
[25] Jha, S. N. (2015). Rapid Detection of Food Adulterants and Contaminants: Theory and Practice. Academic Press.
[26] Jongwanich, J. (2009). The impact of food safety standards on processed food exports from developing countries. *Food Policy*, *34*(5), 447-457.
[27] Jozwiak, A., Milkovics, M., & Lakner, Z. (2016). A network-science support system for food chain safety: a case from Hungarian cattle production. *International Food and Agribusiness Management Review*, *19*(1030-2016-83145), 1-26.
[28] Kępińska-Pacelik, J., & Biel, W. (2021). Microbiological Hazards in Dry Dog Chews and Feeds. *Animals*, *11*(3), 631.
[29] Kleter, G.A., Prandini, A.M., Filippi, L.L., & Marvin, H.J. (2009). Identification of potentially emerging food safety issues by analysis of reports published by the European Community's Rapid Alert System for Food and Feed (RASFF) during a four-year period. Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association, 47 5, 932-50.
[30] Kononiuk, A. D., & Karwowska, M. (2017). Meat and meat products–analysis of the most common threats in the years 2011-2015 in Rapid Alert System for Food and Feed (RASFF). *Roczniki Państwowego Zakładu Higieny*, *68*(3).
[31] Mincer, M., & Niewiadomska-Szynkiewicz, E. (2012). Application of social network analysis to the investigation of interpersonal connections. *Journal of Telecommunications and information technology*, 83-91.
[32] Mukherjee, A., Satija, D., Sinha, S., & Sarma, A. P. (2019). Food Imports in India: Prospects, Issues and Way Forward. *Journal of Economic Science Research*, *2*(3).
[33] Mutukumira, A., & Jukes, D. J. (2003, May). The Development of National Food Safety Control Systems in Sub-Saharan Africa-Issues and Opportunities. In *Food Africa, Congress centre/palais de congress, Yaoundé, Cameroon* (pp. 5-9).
[34] Naughton, D. P., Nepusz, T., & Petroczi, A. (2015). Network analysis: A promising tool for food safety. *Current Opinion in Food Science*, *6*, 44-48.
[35] Nes, K., & Schaefer, K. A. (2018). *Hazard Mitigation in EU Food Trade* (No. 2229-2019-1917).
, and C etc.
[36] Pal, M., Gerbaba,T., Kumar, A. & Shukla,P.K. (2015).Impact of food safety on One Health. Beverage and Food World 42.
[37] Papapanagiotou, E. P. (2017). Foodborne norovirus state of affairs in the EU rapid alert system for food and feed. *Veterinary sciences*, *4*(4), 61.

[38] Papapanagiotou, E. P. (2021). Serious Notifications on Food Contact Materials in the EU RASFF. *Veterinary Sciences*, *8*(4), 56.
[39] Pennone, V., Lehardy, A., Coffey, A., Mcauliffe, O., & Jordan, K. (2018). Diversity of Listeria monocytogenes strains isolated from Agaricus bisporus mushroom production. *Journal of applied microbiology*, *125*(2), 586-595.
[40] Petroczi, A., Nepusz, T., Taylor, G., & Naughton, D. (2011). Network analysis of the RASFF database: a mycotoxin perspective. World Mycotoxin Journal, 4(3), 329-338.
[41] Pigłowski, M. (2017). Product categories and hazard categories in the RASFF notifications: Dependences between chosen variables. *Quality Assurance and Safety of Crops & Foods*, *9*(3), 335-344.
[42] Pigłowski, M. (2017). Notifications to food from European Union countries in the rasff. *Studia i Prace WNEiZ US*, (47/3), 345-358.
[43] Pigłowski, M. (2018). Heavy Metals in notifications of rapid alert system for food and feed. *International journal of environmental research and public health*, *15*(2), 365.
[44] Pigłowski, M. (2019). Pathogenic and Non-Pathogenic Microorganisms in the Rapid Alert System for Food and Feed. *International journal of environmental research and public health*, *16*(3), 477.
[45] Pigłowski, M. (2020). Food hazards on the European Union market: The data analysis of the Rapid Alert System for Food and Feed. *Food science & nutrition*, *8*(3), 1603-1627.
[46] Pigłowski, M. (2021). The Intra-European Union Food Trade with the Relation to the Notifications in the Rapid Alert System for Food and Feed. *International journal of environmental research and public health*, *18*(4), 1623.
[47] Popp, J., Kiss, A., Oláh, J., Máté, D., Bai, A., & Lakner, Z. (2018). Network analysis for the improvement of food safety in the international honey trade. *Amfiteatru Economic*, *20*(47), 84-98.
[48] Robson, K., Dean, M., Haughey, S. A., & Elliott, C. T. (2021). The identification of beef crimes and the creation of a bespoke beef crimes risk assessment tool. *Food Control*, *126*, 107980.
[49] Saurkar, A.V., Pathare, K.G., & Gode, S.A. (2018). An Overview On Web Scraping Techniques And Tools, International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRSCE), Volume 4 Issue 4, PP: 363 – 367.
[50] Schaeffer, S. E. (2007). Graph clustering. *Computer science review*, *1*(1), 27-64.
[51] Smith, M. C., Madec, S., Coton, E., & Hymery, N. (2016). Natural co-occurrence of mycotoxins in foods and feeds and their in vitro combined toxicological effects. *Toxins*, *8*(4), 94.
[52] Somorin, Y. M., Odeyemi, O. A., & Ateba, C. N. (2021). Salmonella is the most common foodborne pathogen in African food exports to the European Union: Analysis of the Rapid Alert System for Food and Feed (1999–2019). *Food Control*, *123*, 107849.
[53] Timmis, H. (2017). The Impact of Standards on Developing Country Exports.
[54] Tudela-Marco, L., Garcia-Alvarez-Coque, J. M., & Martí-Selva, L. (2017). Do EU member states apply food standards uniformly? A look at fruit and vegetable safety notifications. *JCMS: Journal of Common Market Studies*, *55*(2), 387-405.
[55] Van Asselt, E. D., Banach, J. L., & Van Der Fels-Klerx, H. J. (2018). Prioritization of chemical hazards in spices and herbs for European monitoring programs. *Food Control*, *83*, 7-17.
[56] Verbeke, W. (2001). Consumer reactions and economic consequences of the BSE crisis. *Verhandelingen-Koninklijke Academie voor Geneeskunde van Belgie*, *63*(5), 483-492.
[57] Wang, L. (2018). Design Optimization of Food Safety Monitoring System with Social Network Analysis. IEEE Transactions on Computational Social Systems, 5(3), 676-686.
[58] Wu, F., & Guclu, H. (2013). Global maize trade and food security: Implications from a social network model. *Risk analysis*, *33*(12), 2168-2178.
[59] West, D. B. (1996). *Introduction to graph theory* (Vol. 2). Upper Saddle River, NJ: Prentice hall.
[60] Xu, J., Zhao, V. P., Shan, L., & Huang, C. X. (2014). A framework for developing social networks enabling systems to enhance the transparency and visibility of cross-border food supply chains. *GSTF Journal on Computing (JoC)*.